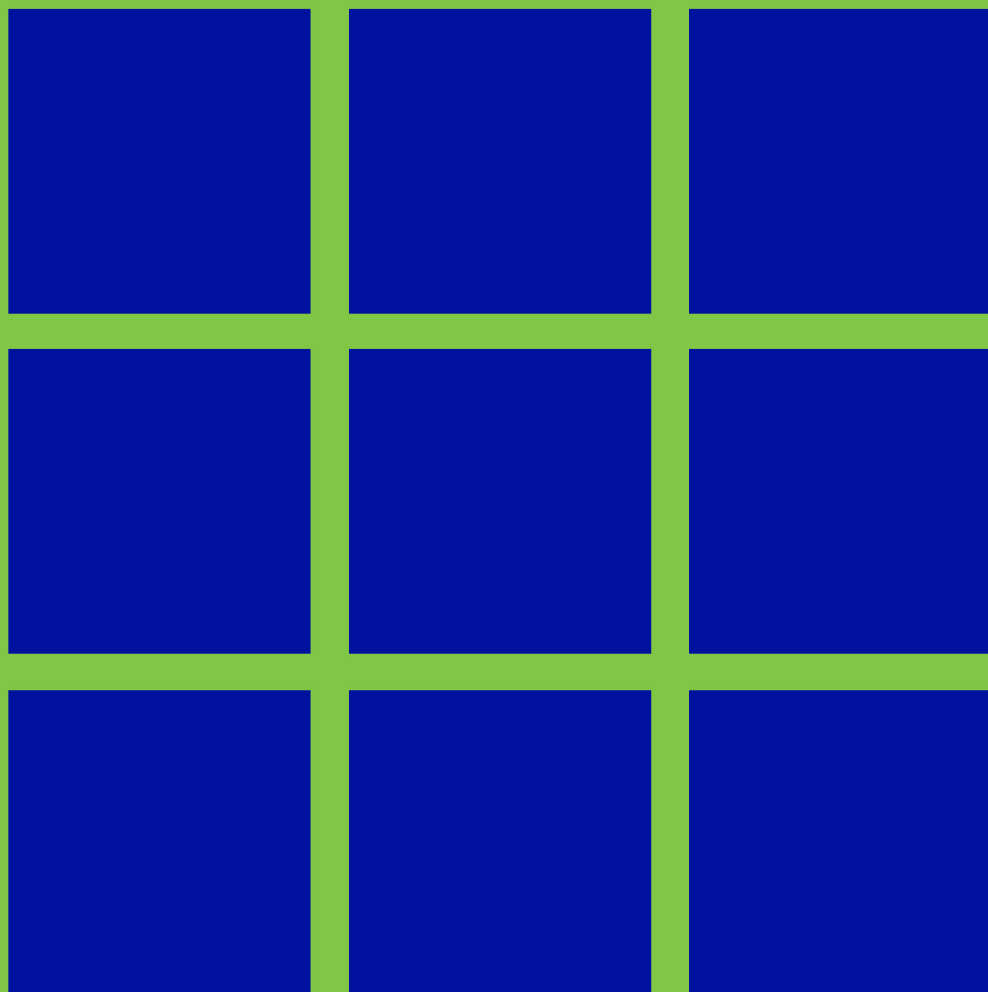


Iolanda Ribeiro & Fernanda Leopoldina Viana (Coord.)



BAL

Bateria de Avaliação da Leitura
abordagens teóricas e opções metodológicas



TESTES PSICOLÓGICOS

BAL

Bateria de Avaliação da Leitura

Abordagens teóricas e opções metodológicas

Coordenação

Iolanda Ribeiro

Fernanda Leopoldina Viana



Testes Psicológicos

CEGOC-TEA, 2014



Obra vencedora do 9.º Prémio CEGOC.

Título

BAL – *Bateria de Avaliação da Leitura*

Coordenadores

Iolanda Ribeiro

Fernanda Leopoldina Viana

Autores

Iolanda Ribeiro

Fernanda Leopoldina Viana

Sandra Santos

Irene Cadime

Séli Chaves-Sousa

Ana Paula Vale

Alina Galvão Spinillo

Consultores

Gerardo Prieto (Universidade de Salamanca, Espanha)

José António Ribeiro Maia (Faculdade de Desporto, Universidade do Porto)

José Junça de Moraes (Universidade Livre de Bruxelas, Bélgica)

Maria Armada Martins da Costa (Faculdade de Letras, Universidade de Lisboa)

Capa

Conceção e *design* da capa by Eduarda Coquet © 2013

Copyright © 2014 by CEGOC-TEA, Lisboa, Portugal.

Primeira edição, Setembro 2014

ISBN: 978-972-8817-75-6

Editor: CEGOC-TEA, Rua General Firmino Miguel, nº3 A/B e Rc, 1600-100 Lisboa.

Edição digital. TODOS OS DIREITOS RESERVADOS.

Proibida a reprodução total ou parcial, sob qualquer forma ou meio, incluindo fotocópia, gravação ou qualquer outro sistema de armazenamento ou recuperação de informação, sem autorização escrita do editor. **ESTA PROIBIÇÃO APLICA-SE A QUALQUER TIPO DE UTILIZAÇÃO, DIVULGAÇÃO OU DIFUSÃO DESTE MATERIAL PARA ALÉM DA UTILIZAÇÃO PESSOAL POR PARTE DO ADQUIRENTE.**

Este trabalho foi desenvolvido no âmbito do projeto "Investigação e avaliação de leitura no 1.º ciclo: Desenvolvimento de inventários de avaliação de leitura" (PTDC/PSI-PED/098592/2008) financiado pela Fundação para a Ciência e Tecnologia, no âmbito do POPH-QREN, participado pelo Fundo Social Europeu e por fundos nacionais do MCTES.

Agradecimentos

Como os leitores poderão constatar ao longo das próximas páginas, um trabalho da envergadura do exigido pela construção da BAL só foi possível graças ao trabalho de uma verdadeira equipa que, ao longo de três anos, nunca se poupou a esforços para que os objetivos fossem plenamente atingidos.

Chegado o momento de partilhar com a comunidade escolar e científica o trabalho produzido, não podemos deixar de registar o nosso profundo reconhecimento aos escritores que connosco colaboraram, produzindo textos originais e apelativos, apesar dos constrangimentos que lhes foram impostos em termos de extensão e de organização. Sem quaisquer tiques de vedetismo e a título gracioso foi possível dar a ler às crianças testadas textos de Ana Maria Magalhães, Carla Maia de Almeida, Isabel Alçada, Isabel Minhós Martins e Jorge Gomes. Em nosso nome, mas também em nome das cerca de 5.000 crianças envolvidas no estudo, o nosso BEM HAJAM!

À nossa colega Eduarda Coquet – que nos respondia sempre “*Sim. Para quando?*” – a todos os pedidos que lhe fazíamos quer estivesse no Porto, quer estivesse no meio do Atlântico, a olhar o azul do cimo de uma colina do Faial, “aquele abraço”.

À Ana Carina Figueiredo, membro do EngageLab da Universidade do Minho e bolseira no projeto *engagebook*, um agradecimento reconhecido pelas aplicações que criou para otimizar um dos testes da BAL, o Teste de Leitura de Palavras, e pela forma “quase mágica” como transformava os nossos pedidos em ferramentas tecnológicas, fazendo-nos crer que era possível fazer (quase) tudo.

Às direções de todos os Agrupamentos Escolares (e foram muitos!) que aceitaram o nosso desafio, assinando um acordo de colaboração com o CIEC – Centro de Investigação em Estudos da Criança, e apoiando toda a logística necessária para que as avaliações se pudessem efetuar com o mínimo de perturbação, impõe-se também uma palavra de agradecimento – Obrigado.

Por fim, a todas as crianças que responderam às sucessivas versões dos testes que integram a BAL e aos seus respetivos professores, reiteramos o nosso agradecimento pelo contributo inestimável que deram à ciência que se faz em Portugal.

As coordenadoras

Iolanda Ribeiro

Fernanda Leopoldina Viana



Autoras

Iolanda Ribeiro

Doutorada em Psicologia, na especialidade de Psicologia da Educação. Professora Auxiliar, Departamento de Psicologia Aplicada da Escola de Psicologia na Universidade do Minho.

Correio eletrónico: iolanda@psi.uminho.pt

Fernanda Leopoldina Viana

Doutorada em Psicologia, na especialidade de Psicologia da Educação. Professora Associada, Departamento de Psicologia da Educação e Educação Especial do Instituto de Educação da Universidade do Minho. Membro integrado do Centro de Investigação em Estudos da Criança – CIEC.

Correio eletrónico: fviana@ie.uminho.pt

Sandra Santos

Licenciada e Mestre em Psicologia Escolar e da Educação. Doutoranda em Psicologia Aplicada, na Escola de Psicologia da Universidade do Minho.

Correio eletrónico: sandra.css@gmail.com

Irene Cadime

Doutorada em Psicologia, na especialidade de Psicologia da Educação. Bolseira de investigação e membro integrado do Centro de Investigação em Estudos da Criança – CIEC.

Correio eletrónico: irenecadime@ie.uminho.pt

Séli Chaves-Sousa

Licenciada e Mestre em Psicologia Escolar e da Educação pela Universidade do Minho. Doutoranda em Psicologia Aplicada, na Escola de Psicologia da Universidade do Minho.

Correio eletrónico: seli.chaves.sousa@gmail.com

Ana Paula Vale

Doutorada em Psicologia. Professora Auxiliar, Departamento de Ciências da Educação e Psicologia da Universidade de Trás-os-Montes e Alto Douro. Membro integrado do Centro de Investigação em Estudos da Criança – CIEC.

Correio eletrónico: pvale@utad.pt

Alina Galvão Spinillo

Doutorada em Psicologia. Professora Catedrática, Departamento de Psicologia, Pós-Graduação em Psicologia Cognitiva da Universidade Federal de Pernambuco, Brasil. Pesquisadora do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Correio eletrónico: alinaspinillo@hotmail.com

ÍNDICE

PREFÁCIO	7
INTRODUÇÃO	11
CAPÍTULO I	15
Avaliar leitura. Questões conceituais e de medida	
<i>Iolanda Ribeiro, Fernanda Leopoldina Viana, Sandra Santos e Irene Cadime</i>	
LER - Questões conceituais	pág. 15
Dificuldades na leitura	pág. 17
A avaliação da leitura	pág. 18
CAPÍTULO II	29
Leitura de palavras. Considerações teóricas e implicações para a sua avaliação	
<i>Fernanda Leopoldina Viana, Ana Paula Vale, Séli Chaves-Sousa e Iolanda Ribeiro</i>	
Modelos teóricos relativos à leitura de palavras	pág. 30
Fatores que influenciam a leitura de palavras	pág. 32
A descodificação: implicações para a sua avaliação	pág. 37
A construção do TLP (Teste de Leitura de Palavras)	pág. 38
CAPÍTULO III	43
Compreensão de textos. Considerações teóricas e implicações para a sua avaliação	
<i>Alina Galvão Spinillo, Iolanda Ribeiro e Fernanda Leopoldina Viana</i>	
Compreensão de textos: considerações teóricas	pág. 43
Compreensão de textos: implicações para a sua avaliação	pág. 48
A construção das provas de compreensão: Testes de Compreensão de Textos na modalidade de leitura (narrativo e informativo) e Testes de Compreensão de Textos na modalidade oral (narrativo e informativo)	pág. 50
CAPÍTULO IV	55
Bateria de Avaliação da Leitura (BAL). Descrição e opções de construção	
<i>Iolanda Ribeiro, Fernanda Leopoldina Viana, Sandra Santos, Irene Cadime, Alina Galvão Spinillo, Séli Chaves-Sousa e Ana Paula Vale</i>	
Plano de construção da BAL	pág. 55
Desenvolvimento dos testes da BAL: opções ao nível da construção e da administração	pág. 57
Materiais que fazem parte da BAL	pág. 60



CAPÍTULO V

63

Opções metodológicas e estatísticas na construção dos testes da Bateria de Avaliação da Leitura

Irene Cadime, Iolanda Ribeiro e Fernanda Leopoldina Viana

Teoria de Resposta ao Item *versus* Teoria Clássica dos Testes | pág. 64

Teoria de Resposta ao Item: o caso específico do modelo Rasch | pág. 66

Estudo das propriedades psicométricas dos testes que integram a BAL | pág. 68

PREFÁCIO

Actualmente, la lectura resulta un instrumento de vital importancia para adquirir conocimiento, para acceder a la cultura, para participar en ella. Este papel tan vital nos debe hacer conscientes de la importancia trascendental que posee la lectura en el mundo educativo, social, cultural y laboral, así como en el propio enriquecimiento personal. La lectura supone un proceso de formación inseparable del proyecto de una sociedad ilustrada por lo que constituye lo que se denomina «cultura lectora», esto es, como habilidades básicas sobre las que se desarrolla toda una actividad cultural, mediante la cual las personas nos desenvolvemos y aplicamos nuestros conocimientos y estrategias lectoras/escritas en múltiples contextos de la vida diaria de manera más o menos eficiente. Aún lo es más desde que interactuamos mediante la tecnología (acceso a Internet desde múltiples artefactos tecnológicos como ordenadores portátiles, tabletas, teléfonos inteligentes,...) desde los cuales podemos leer y comunicarnos de manera instantánea. Tanto es así, que muchos países desarrollados o en vías de desarrollo están dedicando mucho tiempo y esfuerzo a la investigación sobre este tema y también a su evaluación.

Un esfuerzo reciente ha sido el realizado por el *Informe PISA* (Programa Internacional para la Evaluación de Estudiantes) que se viene aplicando desde el año 2000 en todos los países de la OCDE en busca de asumir criterios comunes sobre la evaluación de la competencia lectora. El proyecto PISA de la OCDE es una evaluación internacional estandarizada que se aplica al alumnado de 15 años. En este proyecto, que se inició en el año 2000 y en el que participaron 32 países (y que ha pasado a duplicarse con 65 países participantes en 2009), se evalúa el rendimiento en tres ámbitos: Lectura, Matemáticas y Ciencias. En lo relativo a la lectura, esta prueba lleva a cabo diferentes tipos de tareas lectoras, tales como la elaboración de una comprensión global, la recuperación de información específica, la elaboración de una interpretación o la reflexión sobre el contenido o la forma del texto. Incluye también la habilidad para leer diversos materiales escolares y materiales de lectura no escolares, incluyendo la lectura para el uso personal (cartas personales, ficción, biografías, páginas web, etc.), para el uso público (documentos oficiales, información pública, prensa, etc.) o para el empleo y la educación (libros de texto, informes, etc.).

Además de esta prueba PISA, centrada fundamentalmente en adolescentes de 15 años, existe otra prueba internacional denominada *PIRLS* (Estudio Internacional de Progreso en Comprensión Lectora), y que se viene aplicando a niños de 10 años que cursan sus estudios de primaria en 45 países, al igual que se aplica otra llamada TIMSS, relativa a matemáticas y ciencias. Un aspecto muy revelador de todas estas pruebas es que sus resultados suelen ser coincidentes, esto es, un alumno que obtiene un nivel bajo en la prueba de lectura suele obtener también pobres resultados en las pruebas de matemáticas y ciencias, y al revés, un alumno sobresaliente en las pruebas de lectura suele coincidir con buenas puntuaciones en las pruebas de ciencias.

Todo ello nos incide en la idea de que al igual que ocurre con el aprendizaje, la lectura es un claro ejemplo de procesos cognitivos enormemente complejos que requieren de un largo camino para alcanzar un nivel de competencia adecuado, un nivel de comprensión profundo y eficiente. Hoy sabemos con certeza que los problemas de comprensión que se producen como consecuencia de una inhabilidad lectora, cuyo origen se sitúa frecuentemente en una clara desmotivación del alumno, en edades tempranas y que suele ser el denominador común del fracaso académico en cualquiera de sus fases y desarrollo. Es más, este problema ha trascendido al ámbito extraacadémico, convirtiéndose en un problema social (véase, por ejemplo, el porcentaje tan disparado de paro juvenil en nuestro país que se incrementa más cuanto menor sea la preparación del joven). Por ello, hoy se asume que las dificultades lectoras y de comprensión afecta de manera transversal al alumno y en cuyo estudio están implicadas múltiples áreas de conocimiento. Pero todo ello pasa, necesariamente, por disponer herramientas útiles que precisen, por un lado, un diagnóstico lo más correcto posible y, por otro, una intervención lo más temprana posible.

Como hemos señalado en los párrafos anteriores, hay pruebas lectoras que evalúan el nivel de comprensión y de competencia lectora en niños de 10 años y en adolescentes de 15 años, pero son muy escasas aquellas que se interesan por evaluar el mismo proceso de lectura desde su inicio, desde el momento en el que el niño toma contacto con los libros, con las tareas lectoras en los primeros años de la enseñanza primaria. En esta dirección ha surgido la *Bateria de Avaliação da Leitura*, BAL, consistente en cinco pruebas de lectura, que coordinada por las autoras Iolanda Ribeiro y Fernanda Leopoldina Viana, han desarrollado dicha prueba a la luz de la dirección de diversas investigaciones en las que han participado casi cinco mil niños de edades entre los seis y diez años, junto a su grupo de investigación formado por Alina Galvão Spinillo, Ana Paula Vale, Irene Cadime, Sandra Santos y Séli Chaves-Sousa.

La *Bateria de Avaliação da Leitura* ofrece varias ventajas singulares sobre otras pruebas. Por un lado, esta prueba se sustenta en diversos modelos de lectura que oscilan desde los modelos de lectura de palabras y de comprensión actuales sincronizados con otros modelos y pruebas de lectura y de competencia lectora (PISA, TIMSS, ECOMPLEC), lo que hace fácil y fluida la información que se puede extraer de ella y su conexión con otras pruebas. En otras palabras, esta prueba utiliza un lenguaje en su diagnóstico e interpretación común a otras pruebas muy reconocidas al uso que diagnostican la comprensión lectora en edades posteriores.

Una segunda ventaja es la propia construcción y estructuración de la prueba, muy consistente en su rigor metodológico y estadístico, con valores muy importantes en fiabilidad y validez, y con una muestra muy amplia (cerca de cinco mil participantes distribuidos en cuatro niveles académicos correlativos). Esta batería de lectura adquiere un valor extra al considerar cinco pruebas diferentes construidas en el ámbito de la teoría de respuesta al ítem (modelo de Rasch) y que permite, además, un índice de cada subprueba (tienen un valor en sí mismo). La enorme ventaja que supone esta estructuración de cinco pruebas que pueden ser integradas bajo un mismo paraguas es que permiten un diagnóstico mucho más preciso y certero de las dificultades de lectura detectadas, siendo un valor añadido para la investigación y la práctica profesional. Precisamente, la diversidad de pruebas supone otra de sus ventajas, al ser capaz de distribuir pruebas de lectura oral con escrita (en función de la modalidad de presentación) o de la aplicación de textos de diferente género como pueden

ser los textos narrativos de los informativos. Todo ello supone materiales y actividades consideradas como “naturales” y cotidianas dentro del aula, lo que también facilita su aplicación.

Una cuarta ventaja, ésta especialmente importante, es la de ofrecer una evaluación continua del alumno, esto es, permite un seguimiento de la evolución lectora del alumno en los cuatro primeros años (1º a 4º de primaria) de escolarización, lo que facilita enormemente su diagnóstico y disponer de métodos de intervención. Otro aspecto a tener en cuenta de esta prueba y que supone una nueva ventaja sobre otras pruebas, es su cuidadosa elección de los materiales de lectura, muy apropiados en todos los casos. Las autoras se han esforzado por seleccionar para esta prueba textos adecuados para cada nivel de lectura, texto motivadores en sus contenidos y muy apropiados en su expresividad afectiva en el caso de los textos narrativos. Por último, otra ventaja de esta batería de lectura es la de ofrecer un manual específico de cada subprueba con el objeto de resultar mucho más específico de cada prueba y de esta manera, poder ajustarse mejor a los objetivos de uso de investigación o diagnóstico más particular. Sin duda, este esfuerzo añadido por los autores supone otra ventaja importante de esta prueba. Por otro lado, esta prueba de lectura puede resultar un instrumento muy valioso en la evaluación de la lectura en las edades comprendidas entre los seis y diez años.

José A. León

Catedrático de Psicología de la Universidad Autónoma de Madrid

Septiembre 2014

INTRODUÇÃO

A apresentação dos quadros teóricos que sustentam a construção dos cinco testes da *Bateria de Avaliação da Leitura* – BAL, a descrição do método, especificamente as questões relacionadas com a elaboração dos textos e dos itens das provas de compreensão de textos, bem como as análises estatísticas e as provas usadas nos estudos de validade, são inseridas neste livro e não nos manuais técnicos de cada teste.

Várias razões justificam esta decisão. Em primeiro lugar importa salientar que a bateria de provas foi construída de modo a permitir avaliar a compreensão de textos considerando três condições: o tipo de texto (narrativo ou informativo), a modalidade de apresentação (leitura silenciosa ou audição de um texto) e a leitura de palavras. As questões teóricas que sustentam esta opção são abordadas no primeiro capítulo deste livro. O modelo teórico que enquadra as quatro provas de compreensão é similar, bem como as opções tomadas ao nível da elaboração dos itens e do seu formato. Do ponto de vista metodológico e estatístico, os cinco testes foram construídos no âmbito da Teoria da Resposta ao Item (TRI), mais especificamente com recurso ao modelo Rasch. O plano definido para o estudo da validade dos resultados é igualmente comum aos vários testes. Existe, assim, uma sobreposição de aspetos conceptuais, metodológicos e estatísticos nas provas.

Em segundo lugar, a BAL foi pensada como um conjunto de cinco testes cujos resultados podem ser integrados de modo a permitir um diagnóstico preciso das dificuldades de leitura. Todavia, a BAL pode também ser utilizada em situações de investigação ou de prática profissional. Caberá aos profissionais a opção por utilizar apenas alguns dos testes. Cada teste da BAL foi construído de modo autónomo e cada resultado tem um valor substantivo em si mesmo.

Tendo em conta estas considerações, não se justifica a existência de um único manual técnico relativo a todas as provas. Por um lado, a elaboração de cinco manuais técnicos independentes foi a solução considerada mais ajustada para a utilização dos testes da BAL em função dos objetivos dos seus utilizadores. Por outro lado, dado que o modelo teórico dos testes de compreensão, bem como as opções metodológicas e estatísticas relacionadas com os cinco testes, são comuns, a sua repetição nos cinco manuais resultaria em informação redundante. Optou-se, por isso, pela elaboração de uma publicação complementar aos manuais técnicos, na qual se sistematizam as questões conceptuais e metodológicas que orientaram a construção dos diferentes testes que integram a BAL.

A BAL tem como objetivo avaliar a identificação das palavras escritas (geralmente denominada descodificação), a compreensão da leitura e a compreensão da linguagem oral (embora apresentado num registo elaborado, dado que se trata da audição de um texto que é lido). Este triplo objetivo decorre do reconhecimento dos contributos da habilidade de identificação das palavras escritas e da compreensão da linguagem oral para a compreensão da leitura. Esta bateria responde à necessidade de disponibilizar à

comunidade educativa e científica testes adaptados à população portuguesa a frequentar o 1.º ciclo do ensino Básico e que permitam uma avaliação abrangente.

Consequentemente, a BAL é constituída por cinco testes, com formas ajustadas aos vários anos de escolaridade:

- ◆ *Teste de leitura de palavras* (TLP – 4 formas destinadas aos alunos do 1.º, 2.º, 3.º e 4.º anos);
- ◆ *Teste de compreensão de textos na modalidade oral¹ – narrativo* (TCTMO-n – 4 formas destinadas aos alunos do 1.º, 2.º, 3.º e 4.º anos);
- ◆ *Teste de compreensão de textos na modalidade oral – informativo* (TCTMO-i – 4 formas destinadas aos alunos do 1.º, 2.º, 3.º e 4.º anos);
- ◆ *Teste de compreensão de textos na modalidade de leitura² – narrativo* (TCTML-n – 3 formas destinadas a alunos do 2.º, 3.º e 4.º anos);
- ◆ *Teste de compreensão de textos na modalidade de leitura – informativo* (TCTMO-i – 3 formas, destinadas a alunos do 2.º, 3.º e 4.º anos).

A terminologia adotada nos quatro testes de compreensão traduz a opção por construir um conjunto de provas em que se combina a avaliação da compreensão com a modalidade de apresentação – oral (audição do texto) *versus* leitura (silenciosa) – e com o tipo de texto (narrativo *versus* informativo).

No capítulo I descrevem-se os constructos avaliados pelos testes que constituem a BAL, a par de algumas considerações sobre a avaliação dos mesmos. O capítulo inclui ainda uma secção onde se descrevem as utilizações possíveis da BAL. Um aprofundamento dos constructos “leitura de palavras” e “compreensão” são abordados nos capítulos II e III. No capítulo IV especifica-se o plano de construção da bateria e as características gerais comuns aos testes que a integram. No capítulo V discutem-se as opções metodológicas e estatísticas na construção dos testes da BAL, com especial destaque para as análises do modelo Rasch, um modelo de Teoria de Resposta ao Item. Nos manuais técnicos de cada teste são tratados aspetos específicos relacionados com o processo de construção, instruções para aplicação e cotação e dados psicométricos.

A investigação da leitura exige a utilização de instrumentos com características psicométricas robustas. Neste sentido com a BAL disponibiliza-se aos investigadores um conjunto de testes que permitem uma avaliação rigorosa e abrangente da leitura.

¹ Usaremos a expressão *modalidade oral* quando o examinando escuta a leitura de um texto previamente gravado.

² Usaremos a expressão *modalidade de leitura* quando o examinando procede à leitura silenciosa.

Ao nível da prática, seja no quadro da avaliação, seja no da intervenção junto de alunos com dificuldades de aprendizagem da leitura, os vários testes que constituem a BAL permitirão responder a um conjunto de questões que passamos a enumerar:

- a) Qual o desempenho dos alunos na leitura de palavras na compreensão da leitura e na compreensão de discursos orais, tendo por referência o seu grupo normativo?

As normas disponibilizadas para cada teste por ano de escolaridade permitirão determinar a discrepância entre o desempenho individual e o que é expectável, tendo por referência o respetivo grupo normativo.

- b) Como varia o desempenho na compreensão da linguagem oral e na compreensão da leitura em função do tipo de texto?

As discrepâncias observadas permitirão formular hipóteses sobre a influência desta variável no desempenho dos alunos e, consequentemente, fornecer indicações sobre o trabalho a desenvolver.

- c) Quais as mudanças de desempenho intraindividual que ocorrem ao longo do tempo, quer ao nível da identificação das palavras escritas, quer ao nível da compreensão da leitura?

Como será explicado de modo detalhado no capítulo V, as versões de cada teste para cada um dos anos de escolaridade foram sujeitas a um procedimento de equalização vertical. Esta opção permite que com testes diferentes para cada ano de escolaridade (que atendem às mudanças no desempenho em leitura que decorrem do processo de aprendizagem e com controlo de efeitos de teto) seja possível a avaliação das mudanças intraindividuais e não apenas a comparação dos resultados de um leitor com os do grupo de referência. Desta forma, cria-se a possibilidade de monitorizar os ganhos no tempo e a aproximação ao que é o padrão desejável para determinado ano de escolaridade. Esta possibilidade é particularmente importante para avaliar os efeitos, a curto e a médio prazo, da intervenção junto de alunos com dificuldades na identificação de palavras escritas, na compreensão da leitura e na compreensão da linguagem oral.

- d) Em que medida resultados não satisfatórios ao nível do desempenho em compreensão da leitura variam em função da proficiência na identificação de palavras e/ou na compreensão de enunciados orais?

Esta comparação permitirá formular hipóteses sobre a extensão em que as dificuldades da compreensão da leitura resultam de dificuldades globais de compreensão da linguagem oral ou estão associadas a problemas da descodificação.



Nota final

A identificação das palavras escritas é um dos principais fatores que explica grande parte das dificuldades de compreensão da leitura observada em alunos nos primeiros anos de escolaridade. Porém, apesar de a decodificação ser uma condição necessária para a compreensão da leitura, ela não é suficiente para se chegar a um quadro compreensivo do que pode, de facto, explicar as diferenças interindividuais (Catts, Hogan, & Fey, 2003).

As diferenças interindividuais dependem da compreensão da linguagem oral, mas também de outras variáveis como a capacidade de fazer inferências, da monitorização, da memória de trabalho, do conhecimento lexical e dos conhecimentos prévios do leitor (Perfetti, Landi, & Oakhill, 2005). A avaliação destas variáveis não está contemplada na BAL. Numa perspetiva abrangente de avaliação dos alunos sinalizados com dificuldades de aprendizagem da leitura elas deverão ser incluídas, bem como outras que não se associam exclusivamente a problemas de leitura, como sejam a ocorrência de défices cognitivos ou sensoriais, problemas da esfera emocional e/ou comportamental e características de contextos microssistémicos, em particular os que se relacionam com a família e com o contexto escolar.

Referências bibliográficas

- Catts, H., Hogan, T. P., & Fey, M. (2003). Subgrouping poor readers on the basis of reading-related abilities. *Journal of Learning Disabilities*, 36, 151-164.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227-347). Oxford: Blackwell.

AVALIAR LEITURA. QUESTÕES CONCEPTUAIS E DE MEDIDA

Iolanda Ribeiro, Fernanda Leopoldina Viana, Sandra Santos e Irene Cadime

A construção de provas de avaliação da leitura é uma tarefa complexa dada a natureza do conceito. O desenvolvimento da BAL e a seleção dos constructos a avaliar decorreu, quer de razões de natureza teórica, quer de preocupações relacionadas com a avaliação e a intervenção nos problemas de leitura, em particular dos alunos do 1.º ciclo do Ensino Básico. Da conjugação destes dois fatores resultou a opção de avaliar a leitura de palavras e a compreensão. Ao longo deste capítulo procuramos justificar esta escolha, bem como os procedimentos adotados nos vários testes.

LER - Questões conceptuais

O que é ler?

Entendido como “o processo de simultaneamente extrair e construir significado através da interação e envolvimento com a linguagem escrita” (RAND – Reading Study Group, 2002, p. 11), é consensualmente aceite que ler é um processo que requer a integração de várias fontes de informação. Num sentido menos lato a leitura é, essencialmente, uma habilidade perceptiva e cognitivo-linguística que implica o processamento de informação oriunda de várias fontes. O processamento de informação exige a transformação de uma representação de entrada, geralmente visual (embora possa ser táctil, no caso de leitura pelo sistema Braille) numa outra representação (de saída). De acordo com Morais (2008), este mesmo princípio é aplicado para a compreensão da linguagem oral, ou seja “a representação acústica da fala tem de ser transformada para se obter uma representação do significado da frase, e eventualmente uma representação ortográfica que permite escrevê-la” (p. 1).

Estas operações de transformação – quer se trate de linguagem oral, quer se trate de linguagem escrita – implicam que o leitor ative um conjunto de subprocessos cognitivos (por exemplo, atenção, memória de trabalho) e linguísticos (por exemplo, vocabulário, análise sintática, a capacidade de fazer inferências). No entanto, no caso da leitura, há uma habilidade que lhe é específica – a identificação das palavras escritas. A aquisição desta habilidade é crucial, pois, para que a compreensão seja possível, é preciso o domínio de um código (o código alfabético na maior parte das línguas).

Todavia, o domínio do código é condição necessária mas não suficiente. Ainda de acordo com Morais (2008):

“...para a generalidade dos leitores, a compreensão em leitura (CL) é função da relação multiplicativa entre a compreensão da fala (CF) e a identificação das palavras escritas (IPE). (...). Os casos extremos em que uma das capacidades é muito fraca e a outra é normal também exemplificam esta relação multiplicativa (Gough & Tunmer, 1986). Os *disléxicos*, que apresentam dificuldades importantes em IPE mas não em CF, são afetados na compreensão de textos; e, inversamente, os *hiperléxicos*, que adquiriram os processos necessários à IPE apesar



de terem enormes dificuldades em CF, também o são, neste caso quase ilustrando o fato de que 0 multiplicado por qualquer outro número, por mais elevado que seja, faz 0. Assim, conclui-se que ambos os tipos de conhecimento ou capacidade são necessários e que nenhum deles é suficiente para a compreensão na leitura" (p. 2).

Linguagem oral, decodificação e compreensão da leitura

A aprendizagem da leitura e da escrita são aprendizagens culturais. Nascemos pré-programados para adquirir a linguagem oral, mas não para aprender a ler (e a escrever). A linguagem escrita é uma linguagem de segunda ordem que usa as rotas neuronais da linguagem oral (Dehaene, 2007). Como referem Perfetti, Landi e Oakhill (2005) "*the acquisition of reading comprehension is learning to understand writing as well as one understands spoken language*" (p. 227). Embora existam diferenças importantes entre a linguagem oral, entendida como a capacidade de extrair significados a partir do discurso falado (Snowling & Hulme, 2005), e a linguagem escrita, a investigação tem mostrado de modo sistemático a existência de uma associação entre a compreensão da leitura e a compreensão da linguagem oral (McGuinness, 2005; Nation, 2005), o que indica a possibilidade da existência de processos comuns.

A associação entre a compreensão da leitura e a compreensão da linguagem oral traduz-se em correlações de magnitude variável ao longo do tempo. Na fase inicial da aquisição da leitura, as crianças estão predominantemente orientadas para a aquisição das regras de correspondência grafema-fonema, a que se segue a necessidade de automatizar a leitura de palavras. A compreensão da leitura é limitada por este processo normal de aquisição. Numa primeira fase, as correlações entre os resultados em tarefas de compreensão da leitura são mais elevadas com os da decodificação do que com os da compreensão da linguagem oral (Curtis, 1980; Stanovich, Cunningham, & Freeman, 1984; Sticht & James, 1984).

Numa trajetória de aprendizagem da leitura sem dificuldades associadas à decodificação, à medida que o reconhecimento de palavras se torna cada vez mais automático, reduzem-se as correlações entre a decodificação e a compreensão da leitura e aumentam as correlações entre esta e a compreensão da linguagem oral. Na fase inicial de alfabetização, a leitura de palavras e a compreensão da leitura partilham entre 61 e 81% de variância comum, reduzindo para 45 a 66% a partir do 4.º ano (Bowey, 2000). Em contrapartida, aumenta a percentagem de variância da compreensão da leitura que é explicada pela compreensão da linguagem oral. Estudos com alunos universitários mostram que aquela percentagem se situa em 81% (Gernsbacher, 1990). Mesmo em adultos fluentes mantém-se a associação entre a compreensão da leitura e o reconhecimento de palavras (Perfetti, Landi, & Oakhill, 2005).

Como foi referido, a capacidade para identificar as palavras escritas é específica da leitura. Todavia, sendo específica da leitura e determinante para a extração de sentido, não é suficiente para assegurar a compreensão do que é lido. Identificar uma palavra significa obter a sua forma acústica, mas tal não garante que ative uma representação semântica. Um leitor pode identificar a palavra *sardanapalo* e produzir corretamente a sua forma acústica, mas tal não garante o acesso ao significado, embora ele possa ser inferido através de pistas contextuais caso a palavra esteja integrada numa frase ou num texto.

A compreensão da linguagem oral, a decodificação e a compreensão da leitura têm entre si relações complexas. Esta complexidade decorre de várias razões. Em primeiro lugar, porque considerados isoladamente são, *per se*, constructos complexos. Em segundo lugar porque, durante a aquisição da leitura, se influenciam mutuamente. Em terceiro lugar porque o padrão de interrelação muda ao longo do tempo. Em quarto lugar porque a leitura é também uma competência que muda ao longo da escolaridade, com especial relevo nos anos iniciais (Bowey, 2005). Se estas relações colocam desafios ao nível da investigação no âmbito da leitura, criam exigências particulares quando se trata da construção de provas para a sua avaliação.

Dificuldades na leitura

Dificuldades na leitura de palavras

As dificuldades iniciais na aprendizagem da leitura, no sentido mais estrito da decodificação, apresentam-se associadas a um conjunto de variáveis como o vocabulário, a linguagem recetiva e expressiva, a consciência fonológica, o conhecimento das letras, a nomeação rápida de letras e a compreensão do princípio alfabético (Bowey, 2005; Caravoles, Hulme, & Snowling, 2001).

A sua caracterização tem sido efetuada com base no modelo de dupla via (Coltheart, Rastle, Perry, Langton, & Ziegler, 2001). De acordo com este modelo, a leitura de palavras pode ser feita através de duas vias diferentes: (a) uma via direta, também denominada via lexical ou visual; (b) uma via indireta também denominada subléxica ou fonológica. A utilização de qualquer uma destas vias implica previamente a análise visual da palavra escrita. No entanto, a via lexical implica o acesso direto à representação da palavra armazenada no léxico mental do leitor, ativado pelo contacto com a forma ortográfica da palavra. O acesso ao léxico pela via fonológica faz-se através da conversão grafema-fonema.

Nesta linha é possível distinguir entre subgrupos de dificuldades. Um primeiro subgrupo inclui as crianças nas quais a via fonológica de acesso ao léxico se encontra alterada. Estas crianças conseguem ler palavras regulares ou irregulares, desde que as mesmas sejam familiares, mas apresentam dificuldades na leitura de palavras pouco frequentes ou de pseudopalavras. A dificuldade em utilizar os mecanismos de conversão grafema-fonema explica este padrão de desempenho. Num segundo grupo, os problemas de leitura decorrem da dificuldade em usar a via direta ou léxica. Tipicamente, as crianças deste grupo têm dificuldades em reconhecer uma palavra como um todo e em ler palavras irregulares, sendo o seu desempenho superior na leitura de palavras regulares, sejam estas familiares ou não. Um terceiro grupo inclui as crianças que têm dificuldades na utilização das duas vias de acesso ao léxico (Defior, 2000).

Um leitor proficiente usa as duas vias em função das características das palavras (frequentemente *versus* não frequentes) efetuando uma leitura fluente. No final dos primeiros anos de escolaridade as crianças deverão ser capazes de usar as duas vias de acesso ao léxico e de efetuarem uma leitura automática e precisa das palavras, quer as mesmas se apresentem de modo isolado, quer estejam incluídas num texto. Não atingir o



nível de leitura considerado fluente, ou seja, não ser capaz de efetuar a leitura automática, precisa e expressiva, pode comprometer de forma dramática a compreensão.

As dificuldades na compreensão da leitura

Na comunidade científica, é atualmente aceite que a construção de significados pelo leitor não decorre naturalmente da leitura das palavras presentes num texto (Nation, 2005). A capacidade para identificar as palavras escritas de forma rápida e automática liberta recursos cognitivos para a extração de sentido. Quando o sujeito tem de orientar os seus recursos para processos de nível inferior, i.e., a identificação das palavras escritas, os processos de ordem superior necessários à extração e à construção de significados não são executados ou são-no de modo não eficiente (Bowey, 2005). Assim sendo, não é de estranhar que grande parte das dificuldades ao nível da compreensão se encontre associada a dificuldades ao nível da descodificação (Perfetti, 1985, 1994).

Crianças que apresentam dificuldades na compreensão de textos escritos têm, frequentemente, dificuldades na leitura de palavras; crianças com problemas na leitura de palavras têm, sistematicamente, dificuldades na compreensão de textos escritos (Vellutino & Fletcher, 2005). Estudos realizados por Oakhill (1994) e Yuill e Oakhill (1991) contribuíram para distinguir, no âmbito das dificuldades de leitura, um grupo de alunos que se caracteriza por apresentar na leitura de palavras um desempenho adequado à sua faixa etária, mas com desempenhos em compreensão da leitura abaixo do esperado para essa mesma faixa etária. Este grupo, identificado como “*poor comprehenders*” (Nation, 2005) é, tipicamente, menos referenciado para apoio do que o de alunos com problemas ao nível da leitura de palavras (Leach, Scarborough, & Rescorla, 2003). Excluídos os problemas de leitura de palavras como causa das dificuldades de compreensão da leitura, os resultados de alguns estudos (e.g. Nation & Snowling, 1999) mostram que estes alunos (“*poor comprehenders*”), por comparação com um grupo de controlo sem dificuldades, apresentam um desempenho inferior ao dos seus pares em tarefas de compreensão da linguagem oral (*listening comprehension*). Estas dificuldades poderão indiciar dificuldades mais gerais de compreensão que não se restringem apenas a uma modalidade (oral *versus* leitura). Além de problemas de compreensão da linguagem oral, variáveis como conhecimento lexical e memória de trabalho (Perfetti *et al.*, 1996), realização de inferências e monitorização da compreensão (Nation *et al.*, 2004) estão igualmente presentes nos alunos deste subgrupo.

A avaliação da leitura

Como foi referido anteriormente, a avaliação da leitura deverá incluir: a competência na identificação das palavras escritas (vulgarmente chamada descodificação, embora, em rigor, a descodificação seja, na realidade, a forma mais usada e mais eficaz para identificar as palavras escritas)¹, a compreensão da linguagem oral e a compreensão da leitura. Avaliar apenas a compreensão da leitura não nos permite saber

¹ Nesta obra usaremos *descodificação* e *identificação* de palavras escritas como sinónimos.

o contributo de cada um dos fatores (descodificação e compreensão oral), nem saber as origens de eventuais dificuldades. Consequentemente, *“não é possível, portanto, apurar, sem uma avaliação separada, independente, dos dois tipos de capacidade (compreensão da fala e identificação das palavras escritas), se um aluno ou uma aluna compreende mal os textos porque compreende mal, de modo geral, a linguagem, ou porque tem dificuldades em identificar as palavras escritas”* (Morais, 2008, p. 3).

Se ler é um processo complexo, para o qual concorrem múltiplos fatores e processos cognitivos e linguísticos, facilmente se depreenderá que avaliar a leitura é, igualmente, um processo complexo e que depende também dos objetivos subjacentes a essa avaliação. Se o objetivo é avaliar as mudanças de conhecimentos que se operam depois da leitura, está-se a avaliar um produto. Se o objetivo é avaliar o conjunto de competências que o leitor mobiliza (ou não) ao longo da leitura para que se produzam as mudanças ao nível do conhecimento, estamos a avaliar processos (Viana, 2009).

Estes dois tipos de avaliação não são mutuamente exclusivos, mas complementares. Numa mesma prova podem ser avaliados produtos e processos, dependendo do tipo de análise que for feita e das tarefas nela incluídas. A BAL foi desenvolvida de forma a permitir esta complementaridade.

A fim de explicitar as opções nos formatos adotados nos vários testes da BAL, apresenta-se, de seguida, uma síntese das principais tarefas de avaliação da leitura na vertente descodificação e compreensão.

A avaliação da leitura de palavras – descodificação

A leitura de palavras constitui um objetivo de aprendizagem nos primeiros anos de aquisição da leitura, relativamente ao qual muitas crianças experienciam dificuldades. Estas dificuldades limitam severamente a capacidade de compreensão da leitura (Perfetti, Landi, & Oakill, 2005). A avaliação da leitura de palavras tem, assim, de ser necessariamente contemplada na avaliação da leitura, quer porque em si mesma é uma dificuldade, quer pelo impacto que tem na compreensão da leitura.

A avaliação da leitura de palavras (também denominada avaliação da descodificação ou do reconhecimento de palavras) recorre, maioritariamente, a dois tipos de procedimentos: a leitura de palavras apresentadas de forma isolada e a leitura de palavras apresentadas em (con)texto.

a) Leitura de palavras apresentadas de forma isolada

Neste tipo de tarefa o leitor tem de identificar um conjunto de palavras apresentadas de forma isolada em cartões, no ecrã de um computador ou em listas. Este tipo de tarefa fornece uma medida relativamente pura da habilidade de descodificação, já que o leitor não pode usar pistas contextuais. Todavia, pode sempre socorrer-se de outras estratégias, que serão aprofundadas no capítulo seguinte, tais como fazer uma leitura visual (usando pistas relativas à forma da palavra), ou ler por analogia (Ehri, 1998).

Nos vários instrumentos de avaliação da leitura de palavras apresentadas de forma isolada, podem encontrar-se algumas variantes, relativamente ao formato de apresentação, tais como: i) Ler, num minuto, o número máximo de palavras apresentadas numa lista (por exemplo, *Dyslexia Screening Test*, Nicolson & Fawcett, 1996); ii) Ler, em silêncio, uma palavra e marcar a imagem relativa a essa palavra selecionando a respetiva imagem de entre 3 ou 4 imagens dadas (por exemplo, *Group Reading Assessment and Diagnostic Evaluation*, Williams, 2002; Prova de Reconhecimento de Palavras, Viana & Ribeiro, 2010); iii) Marcar a palavra pronunciada pelo examinador, selecionando-a de entre 3 ou 4 palavras escritas (por exemplo, *Group Reading Assessment and Diagnostic Evaluation*, Williams, 2002).

Em geral, este tipo de tarefa não é apresentado com tempo limite, embora se possam encontrar exceções, tal como a referida acima e que integra o *Dyslexia Screening Test* ou a Prova de Reconhecimento de Palavras. A ordem de apresentação das palavras pode ser aleatória ou hierarquizada por grau de dificuldade.

As palavras a apresentar de forma isolada para avaliar a decodificação são, geralmente, selecionadas de acordo com um conjunto de critérios psicolinguísticos. Os mais usados são: (i) a frequência, (ii) a regularidade, e (iii) a extensão.

Alguns testes incluem também pseudopalavras. O uso de pseudopalavras para avaliar a decodificação não é consensual. Se pode, por um lado, ser considerada a medida mais pura para testar o processamento fonológico, por outro lado há estudos que reportam diferenças de desempenho em função do tipo de pseudopalavras usado. Quando as pseudopalavras são itens similares, em termos visuais, a palavras reais, podem induzir o uso da leitura por analogia, reduzindo a exigência quanto à competência em avaliação – o processamento fonológico (Rathvon, 2004).

b) Leitura de palavras em (con)texto

Neste tipo de tarefa o leitor é convidado a ler um texto, sendo registada a precisão da leitura efetuada. Trata-se de um procedimento muito utilizado em testes de fluência, nos quais, além da precisão, é avaliada também a velocidade de leitura (número de palavras corretamente lidas por minuto) e, embora mais raramente, a prosódia. Tem a vantagem de confrontar o leitor com uma tarefa mais usual – a leitura de um texto – mas, em contrapartida, o leitor, mesmo o menos fluente, pode fazer uso de pistas contextuais para a identificação das palavras. Na leitura de palavras em (con)texto, o controlo das palavras a usar, do ponto de vista das variáveis psicolinguísticas atrás referidas, é bastante mais difícil de conseguir.

Na BAL optou-se por integrar um teste de leitura de palavras, selecionadas tendo em conta os critérios psicolinguísticos atrás referidos e ainda o tipo de estruturas silábicas. Quanto a este último critério, incluíram-se as estruturas silábicas mais frequentes em Português Europeu (Viana, Ribeiro, Vale, Chaves-Sousa, Santos, & Cadime, 2014). A sua construção tentou dar resposta a um desafio adicional – a produção de formas diferentes do *Teste de Leitura de Palavras* por ano de escolaridade. Uma única lista de palavras, a ser aplicada sem tempo limite em diferentes anos de escolaridade, teria de ser uma lista extensa, com itens que seriam forçosamente muito difíceis para os anos escolares iniciais e muito fáceis para os anos escolares mais

avançados. Segundo as recomendações formuladas no documento *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999), foram elaboradas 4 listas calibradas através do uso de procedimentos de equalização para assegurar a comparabilidade entre resultados entre as diferentes formas do teste. Este procedimento – denominado equalização vertical – tem a vantagem de permitir a comparação de resultados em habilidades que mudam ao longo do tempo, com especial relevância para os anos iniciais de escolaridade, como é o caso da identificação de palavras (descodificação) e da compreensão em leitura. A descrição detalhada do processo de calibração dos testes que integram a BAL será apresentada no capítulo V.

A avaliação da compreensão da leitura

A variável texto

Embora em alguns estudos a variável texto (narrativo *versus* informativo) não se tenha revelado diferenciadora em termos de desempenho em leitura, a maior parte da investigação aponta esta variável como sendo determinante para explicar diferenças ao nível da compreensão (Giasson, 2005; Sim-Sim, Duarte, & Micaelo, 2007).

Uma das razões explicativas para esta diferença poderá ser a maior proximidade dos textos narrativos ao discurso oral. De acordo com Stein e Albro (1997), a compreensão de uma história depende em larga medida do conhecimento acerca das intenções que motivam a ação humana e este conhecimento é adquirido precocemente. Segundo estes autores, aos três anos, este conhecimento já está presente nas crianças. Uma razão adicional pode ser aduzida: quer nos anos pré-escolares, quer no 1.º ciclo do Ensino Básico, a leitura “pela voz dos outros” e a leitura autónoma são efetuadas maioritariamente com recurso a obras de literatura para a infância, quase exclusivamente narrativas. No TCTMO-n e no TCTML-n, são usados textos narrativos e, no TCTMO-i e no TCTML-i, textos predominantemente informativos².

As tarefas

No que respeita às tarefas para a avaliação da compreensão da leitura, as perguntas ocupam o lugar cimeiro (perguntas de verdadeiro/falso, perguntas de resposta de escolha múltipla e perguntas de resposta aberta), seguidas de perto pelos procedimentos de completamento (*cloze*) (Viana, 2009).

a) Perguntas do tipo verdadeiro/falso

Não requerem uma resposta verbal complexa (Cain & Oakhill, 2006), o que torna este tipo de tarefa particularmente útil para avaliar crianças com dificuldades na expressão escrita ou oral (Greene, 2001; Viana,

² Quando os textos relatam temas científicos destinados a crianças, geralmente adota-se uma estratégia de narrar para explicar, recorrendo a textos que são simultaneamente informativos e narrativos. Os textos informativos usados são, de facto, textos de natureza híbrida, nos quais a organização narrativa funciona como uma espécie de pano de fundo que suporta a sequência informativa, captando a atenção e estimulando a curiosidade das crianças. O espaço, o tempo e as personagens constroem a estrutura que apoia a explicação e a descrição científica.

2009). Têm, todavia, limitações na avaliação de competências relevantes para a compreensão da leitura, nomeadamente quanto à capacidade de discriminar diferentes níveis de competências na realização de inferências (Cain & Oakhill, 2006; Oakhill, Yuill, & Parkin, 1986), bem como limitações que podem resultar da elevada probabilidade (50%) de acertos com respostas ao acaso.

b) Perguntas com resposta de escolha múltipla

Têm como vantagem a rapidez nos procedimentos de avaliação e de cotação, o que permite incluir um número elevado de perguntas num único teste e, deste modo, aumentar o espectro de níveis de processamento de informação incluídos (Campbell, 2005). Este formato de perguntas permite ainda excluir a influência das competências de escrita, pelo que tem sido amplamente adotado em testes standardizados de avaliação da compreensão da leitura (Snow, 2003). As principais críticas a este tipo de perguntas dizem respeito às limitações relativas à avaliação das competências para realizar inferências e das estratégias de leitura que ocorrem durante a leitura (Carver, 1992), ao facto de os examinados poderem responder com relativa precisão às perguntas sem terem lido os textos (Daneman & Hannon, 2001; Keenan & Betjemann, 2006) e ao facto de apenas permitir verificar se o leitor reconhece os significados que os autores das perguntas previamente construíram (Valencia & Pearson, 1987). Reconhecendo vantagens e limitações em todos os formatos de tarefa, consideramos que as desvantagens podem ser contornadas, através de um rigoroso controlo da formulação das perguntas e das respetivas alternativas de resposta.

c) Perguntas de resposta aberta

Após a leitura de um texto, os leitores respondem, oralmente ou por escrito, a um conjunto de perguntas (Cain & Oakhill, 1996). Uma das vantagens deste formato de pergunta prende-se com o seu potencial para eliciar envolvimento com o texto (Campbell, 1999), para além de a análise das respostas erradas poder ajudar à análise operativa do processo de aprendizagem, permitindo clarificar o que esteve na base do erro (Cain, Oakhill, Barnes, & Bryant, 2001; Cain, Oakhill, & Elbro, 2003). Como desvantagem, é apontado o facto de os resultados obtidos poderem ser influenciados por competências que não são tradicionalmente classificadas como componentes da compreensão, nomeadamente as competências declarativas e de escrita (Jenkins, Johnson, & Hileman, 2004; Shanahan, 1984; Shanahan & Lomax, 1986).

d) Tarefas do tipo cloze

Nestas tarefas os leitores são convidados a preencher, com as palavras contextualmente adequadas, espaços em branco existente num conjunto de frases. As palavras a usar podem ser livremente escolhidas pelo leitor, ou seleccionadas de entre uma lista de alternativas (geralmente não superior a cinco) que lhe é fornecida por cada espaço a preencher (Cain & Oakhill, 2006). Dado que este tipo de tarefa exige que o leitor estabeleça relações entre os elementos do texto e as associe e integre nos seus conhecimentos prévios, tem a vantagem de reduzir potenciais fontes de viés que outros formatos de resposta podem apresentar como, por exemplo, uma má formulação de perguntas (Greene, 2001). Como principais desvantagens regista-se o facto de o desempenho do leitor estar muito dependente dos conhecimentos prévios possuídos sobre o tema (Cohen, 1975; Page,

1975) e do tipo de palavra requerido para o preenchimento da lacuna (adjetivo, verbo, substantivo...) (Abraham & Chapelle, 1992).

Atendendo aos objetivos da BAL, a opção recaiu no formato de perguntas com resposta de escolha múltipla, quer devido à facilidade de cotação, especialmente útil para que possa ser utilizada em avaliações com grande número de sujeitos, quer porque se pretendia que viesse a ser usada na avaliação da compreensão nas duas modalidades (oral *vs* escrita). O uso da técnica de *cloze* e de perguntas de resposta aberta criaria dificuldades adicionais de aplicação e cotação na prova de compreensão de textos na modalidade oral.

A avaliação da compreensão da linguagem falada

Linguagem oral e linguagem escrita diferem em vários aspetos. A linguagem oral é contextualizada, o interlocutor está presente e o recetor pode pedir ao emissor que clarifique o que não entendeu. Além disso, o registo escrito não é linguagem oral transposta para escrita, mas uma codificação numa outra linguagem. Todavia, as teorias da compreensão não procedem a uma distinção clara entre compreensão da linguagem oral e compreensão da linguagem escrita (Cain & Oakhill, 2008), considerando que o produto é idêntico – uma representação mental coerente, integrada e completa (Johnson-Laird, 1993; Kintsch, 1998) e que os processos para construir esta representação mental são os mesmos, independentemente da modalidade de apresentação – escrita ou falada (Gernsbacher, 1990).

Analisando os resultados da investigação sobre leitura de palavras, compreensão da leitura e compreensão da linguagem oral, dois aspetos principais são de salientar. Em primeiro lugar a compreensão da leitura e a compreensão da linguagem oral estão interrelacionadas e influenciam-se mutuamente. Se a compreensão da linguagem oral se constitui como um fator de explicação das diferenças interindividuais e das dificuldades de compreensão da leitura, esta, por sua vez, influencia o desenvolvimento da capacidade de compreender a linguagem oral (Perfetti, Landi, & Oakill, 2005). Assim, embora os processos para construir representações mentais a partir de produções linguísticas – lidas ou ouvidas – sejam os mesmos, num quadro compreensivo de avaliação da leitura e das suas dificuldades, torna-se necessário avaliar a compreensão da leitura e a compreensão da linguagem oral, dado que existem sujeitos que têm desempenhos diferenciados nas duas áreas (Nation, 2005).

Referências bibliográficas

- Abraham, R. G., & Chapelle, C. A. (1992). The meaning of cloze test scores: An item difficulty perspective. *The Modern Language Journal*, 76(4), 468-479.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

- Bowey, J. A. (2000). Recent developments in language acquisition and reading research: The phonological basis of children's reading difficulties. *Australian Educational and Developmental Psychologist*, 17, 5-31.
- Bowey, J. A. (2005). Predicting individual differences in learning to read. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 155-172). Oxford: Blackwell.
- Cain, K. & Oakhill, J. (1996). The nature of the relationship between comprehension skill and the ability to tell a story. *British Journal of Developmental Psychology*, 14(2), 187-201.
- Cain, K., & Oakhill, J. (2006). Assessment matters: Issues in the measurement of reading comprehension. *British Journal of Educational Psychology*, 76, 697-708.
- Cain, K. & Oakhill, J. (2008). Cognitive bases of children's language comprehension difficulties. Where do we go from here? In K. Cain & J. Oakhill (Eds.), *Children's comprehension problems in oral and written language. A cognitive perspective* (pp. 283-295). New York: The Guilford Press.
- Cain, K., Oakhill, J., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, 29(6), 850-859.
- Cain, K., Oakhill, J., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language*, 30(3), 681-694.
- Campbell, J. R. (1999). *Cognitive processes elicited by multiple-choice and constructed response questions on an assessment of reading comprehension* (Unpublished doctoral dissertation). Temple University, Philadelphia, USA.
- Campbell, J. R. (2005). Single instrument, multiple measures: Considering the use of multiple item formats to assess reading comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 347-368). New Jersey: Lawrence Erlbaum Associates Publishers.
- Caravolles, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language*, 45, 751-774.
- Carver, R. P. (1992). What do standardized tests of reading comprehension measure in terms of efficiency, accuracy, and rate? *Reading Research Quarterly*, 27(4), 346-359.
- Cohen, J. H. (1975). The effect of content area material on Cloze test performance. *Journal of Reading*, 19(3), 247-250.
- Coltheart, M., Rastle, K., Perry, C., Langton, R., & Ziegler, J. (2001). DRC: A dual-route cascade model of visual word recognition and reading aloud. *Psychological Review*, 108, 204-256.
- Curtis, M. E. (1980). Vocabulary testing and vocabulary instruction. In M. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 37-51). Hillsdale, NJ: Erlbaum.
- Daneman, M., & Hannon, B. (2001). Using working memory theory to investigate the construct validity of multiple-choice reading comprehension tests such as the SAT. *Journal of Experimental Psychology: General*, 130(2), 208-223.

- Defior, S. (2000). *Dificuldades de aprendizagem: Um enfoque cognitivo. Lectura, Escritura, Matemáticas*. Granada: Ediciones Aljibe.
- Dehaene, S. (2007). *Les neurones de la lecture*. Paris: Odile Jacob.
- Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. *In* J. Metsala & L. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 3-40). Mahwah, NJ: Erlbaum.
- Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.
- Giasson, J. (2005). *La lecture: De la théorie à la pratique*. Bruxelles: De Boeck.
- Gough, P. & Tunmer, W. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, 7, 6-10.
- Greene, B. (2001). Testing reading comprehension of theoretical discourse with cloze. *Journal of Research in Reading*, 24(1), 82-98.
- Jenkins, J. R., Johnson, E., & Hileman, J. (2004). When is reading also writing: Sources of individual differences on the new reading performance assessments. *Scientific Studies of Reading*, 8(2), 125 - 151.
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, UK: Cambridge University Press.
- Keenan, J. M. & Betjemann, R. S. (2006). Comprehending the Gray Oral Reading Test without reading it: Why comprehension tests should not include passage-independent items. *Scientific Studies of Reading*, 10(4), 363-380.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emergent reading disabilities. *Journal of Educational Psychology*, 95, 211-224.
- McGuinness, D. (2005). *Language development and learning to read. The scientific study of how language development affects reading skill*. Cambridge, Mass.: The MIT Press.
- Morais, J. (2008). A aprendizagem da leitura: Condições, capacidades envolvidas e trajetórias. Conferência proferida no *Seminário Internacional Profissão Professor: o resgate da pedagogia*. Instituto Alfa e Beto, agosto-setembro 2008. Acedido em 3 de junho de 2013:
http://alfaebeto.org.br/profissaoprofessor/administrator/pdf/artigo_seminario_2009_jose_morais.pdf
- Nation, K. (2005). Children's reading comprehension difficulties. *In* M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 248-265). Oxford: Blackwell Publishers.
- Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment. *Journal of Speech, Hearing and Language Research*, 41(1), 199-211.
- Nation, K. & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: evidence from semantic priming. *Cognition*, 70, B1-B13.

- Nicolson, R. I., & Fawcett, A. J. (1996). *Dyslexia Early Screening Test*. London: Psychological Corporation.
- Oakhill, J. V. (1994). Individual differences in children's text comprehension. *In* M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 821-848). San Diego, CA: Academic.
- Oakhill, J., Yuill, N., & Parkin, A. (1986). On the nature of the difference between skilled and less-skilled comprehenders. *Journal of Research in Reading*, 9(2), 80-91.
- Page, W. D. (1975). The post oral reading cloze test: New link between oral reading and comprehension. *Journal of Reading Behavior*, 7(4), 383-389.
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- Perfetti, C. A. (1994). Psycholinguistics and reading ability. *In* M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 849-894). San Diego, CA: Academic.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. *In* M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227-347). Oxford: Blackwell.
- Perfetti, C. A., Marron, M. A., & Foltz, P. W. (1996). Sources of comprehension failure: Theoretical perspectives and case studies. *In* C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 137-165). Mahwah, NJ: Erlbaum.
- RAND Reading Study Group (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND Corporation.
- Rathvon, N. (2004). *Early reading assessment: A practitioner's handbook*. New York: The Guilford Press.
- Shanahan, T. (1984). Nature of the reading-writing relation: An exploratory multivariate analysis. *Journal of Educational Psychology*, 76(3), 466-477.
- Shanahan, T., & Lomax, R. G. (1986). An analysis and comparison of theoretical models of the reading-writing relationship. *Journal of Educational Psychology*, 78(2), 116-123.
- Sim-Sim, I., Duarte, C., & Micaelo, M. (2007). *O ensino da leitura: A compreensão de textos*. Lisboa: Direcção-Geral de Inovação e Desenvolvimento Curricular.
- Snow, C. E. (2003). Assessment of reading comprehension: Researchers and practitioners helping themselves and each other. *In* A. P. Sweet & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 192-206). New York: The Guilford Press.
- Snowling, M. & Hulme, C. (Eds.) (2005). *The science of reading: A handbook*. Oxford: Blackwell Publishers.
- Stanovich, K. E., Cunningham, A. E., & Freeman, D. J. (1984). Relation between early reading acquisition and word decoding with and without context: A longitudinal study of first grade children. *Journal of Educational Psychology*, 76, 668-677.

- Stein, N. L. & Albro, E. R. (1997) Building complexity and coherence: Children's use of goal-structured knowledge in telling stories. In M. Bamberg (Ed.), *Narrative development: Six approaches* (pp. 5-44). Mahwah, NJ: Lawrence Erlbaum.
- Sticht, T. & James, J. (1984). Listening and reading. In P. Pearson (Ed.), *Handbook of research on reading* (pp. 293-317). New York: Longman.
- Valencia, S. & Pearson, P. (1987). Reading assessment: Time for a change. *The Reading Teacher*, 40, 726-732.
- Vellutino, F. R. & Fletcher, J. N. (2005). Developmental dyslexia. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 362-378). Oxford: Blackwell.
- Viana, F. L. (2009). *O ensino da leitura: a avaliação*. Lisboa: Direção Geral de Inovação e Desenvolvimento Curricular.
- Viana, F. L. & Ribeiro, I. (2010). *PRP - Prova de Reconhecimento de Palavras*. Lisboa: CEGOC-TEA.
- Viana, F. L., Ribeiro, I., Vale, A. P., Chaves-Sousa, S., Santos, S., & Cadime, I. (2014). *TLP - Teste de leitura de palavras. Manual técnico*. Lisboa: CEGOC-TEA.
- Williams, K. T. (2002). *Group reading assessment and diagnostic evaluation*. Circle Pines, MN: American Guidance Service.
- Yuill, N. & Oakhill, J. (1991). *Children's problems in text comprehension*. Cambridge: Cambridge University Press.

LEITURA DE PALAVRAS. CONSIDERAÇÕES TEÓRICAS E IMPLICAÇÕES PARA A SUA AVALIAÇÃO

Fernanda Leopoldina Viana, Ana Paula Vale, Séli Chaves-Sousa e Iolanda Ribeiro

A tarefa principal de uma criança que inicia a aprendizagem da leitura é entender como a escrita representa a língua que já conhece. A ortografia portuguesa usa 23 letras (às quais se juntaram recentemente o <k>, o <y> e o <w>), com as quais é possível representar os sons da língua falada. Como se depreende, para poder representar sons através de letras, é necessária a identificação dos sons, ou seja, é necessária a consciência fonémica. Qualquer que seja o método de iniciação à leitura, ele tem de ajudar o aprendiz a aceder ao fonema, pois são os sons e não as letras a base do código escrito (McGuinness, 2006). Aprender a ler num sistema de escrita alfabética exige o domínio de um código (no nosso caso, o código alfabético). Dominar este código é a habilidade específica da leitura. Dificuldades a este nível comprometem inexoravelmente a extração de sentido (Snowling & Hulme, 2005).

A descodificação, como o próprio nome indica, é a habilidade de aplicar um código para decifrar uma mensagem. Usado na leitura, o conceito de descodificação refere-se ao processo que leva à recuperação da fonologia, ou seja da pronúncia de um padrão ortográfico que é, em geral, uma palavra. Descodificar é o aspeto técnico da leitura, a aplicação do princípio alfabético¹. Para que as crianças aprendam a descodificar com eficácia é necessário um ensino explícito sobre as relações letra (grafema) - som (fonema) (Morais, 1997). Através da prática, a descodificação progride de um exercício laborioso e lento de conversão grafema-fonema, passando pela apropriação das regularidades que ocorrem em estruturas que são partes de palavra, para se tornar num modo de reconhecimento rápido e automático de padrões específicos de palavras (Ehri, 2005).

Como foi referido no capítulo anterior, a aprendizagem da leitura é complexa e pode tornar-se muito difícil para algumas crianças. As crianças que apresentam dificuldades em aprender a ler manifestam-nas primariamente ao nível da descodificação (Høien-Tengesdal & Tønnessen, 2011; Snowling & Hulme, 2012).

Neste capítulo, abordaremos de forma sucinta os modelos teóricos relativos à identificação das palavras escritas e aos fatores que a influenciam, de modo a contextualizar o leitor quanto às opções que sustentam, quer a inclusão, na BAL, de um teste de leitura de palavras, quer a sua elaboração.

¹ Princípio alfabético – pressuposição que fundamenta os sistemas de escrita alfabéticos, segundo a qual cada som da fala ou fonema de uma língua deve ter a sua própria representação gráfica e distinta (Harris & Hodges, 1999, p. 217).

Modelos teóricos relativos à leitura de palavras

A investigação sobre os processos subjacentes à leitura (e.g. Ellis, 1989; Goodman, 1968; Gough, 1972; LaBerge & Samuels, 1974; Morais, 1994; Rumelhart, 1977, 1994; Samuels, 1994; Smith, 1973; Stanovich, 1980) foi gerando uma série de teorias e de modelos de leitura. Assim, são numerosos os estudos efetuados a partir dos anos 70 do século XX, centrados na análise das operações e estratégias que estão presentes no ato de ler.

As diferentes conceptualizações sobre a aprendizagem da leitura podem ser organizadas em três grandes grupos de modelos: i) os modelos ascendentes; b) os modelos descendentes e; iii) os modelos interativos.

Os *modelos ascendentes* (e.g. o modelo de Gough, 1972, ou o modelo de LaBerge e Samuels, 1974) como o próprio nome indica, consideram que a leitura é um percurso linear e hierarquizado que parte dos processos psicológicos mais básicos – identificar e juntar letras – para chegar aos processos psicológicos de ordem superior – a extração de sentido. Perante um texto escrito o leitor começa por identificar as letras, combiná-las em sílabas que, por sua vez, se reúnem para formar palavras. As palavras, por sua vez, integram frases que, combinadas resultam em textos. A linguagem escrita é, para estes modelos, a codificação da linguagem oral, e ler a capacidade de descodificar ou traduzir a mensagem escrita no seu equivalente oral.

Os *modelos descendentes* (e.g. o modelo Goodman, 1968, ou o modelo de Smith, 1973) consideram que a leitura segue um processo completamente inverso ao anterior. A leitura é concebida como “um jogo de adivinhas psicolinguísticas”, no qual o leitor, utilizando os seus conhecimentos prévios sobre o tema e o contexto, faz antecipações que são confirmadas ou infirmadas pelos índices fornecidos pelo texto escrito. Os processos de ordem superior são primeiramente ativados, através de antecipações, pelo que a leitura é um processo visuoperceptivo de reconhecimento de palavras (sem descodificação).

As críticas a estes modelos, nomeadamente quanto às suas insuficiências para explicar muitas situações com que o leitor se depara, conduziram ao aparecimento de modelos dinâmicos – os modelos interativos (e.g. Ellis, 1989; Rumelhart, 1977 ou Stanovich, 1980) – os quais consideram que a leitura exige, em simultâneo e em interação, processos de ordem inferior (como processos perceptivos e lexicais) e processos de ordem superior.

Assim, para os *modelos interativos*, ao ler uma palavra, o leitor ativa uma via direta de acesso ao significado (também chamada visual ou lexical) se essa palavra for muito familiar, ou uma via indireta (também chamada fonológica ou sublexical), se essa palavra for desconhecida.

O reconhecimento imediato de uma palavra escrita depende muito da experiência de leitura. Embora alguns métodos de iniciação à leitura partam da memorização de um conjunto de palavras, a partir das quais, através de uma abordagem descendente, se chega ao grafema, o uso eficiente da via direta só é possível com experiência de leitura. No início da escolaridade, mesmo que o aprendiz leitor possua um léxico alargado, ele é essencialmente fonológico (pois conhece as palavras ditas e ouvidas) e não ortográfico (ou seja, não conhece a esmagadora maioria das formas escritas dessas mesmas palavras).

A experiência de leitura permite ainda que unidades ortográficas mais amplas do que a letra (como, por exemplo dígrafos, tais como <rr> ou <nh>, ou estruturas silábicas muito frequentes na língua) sejam processadas como padrões e não como uma sucessão de letras. Morais (1997) propõe um modelo de leitura de palavras (cf. figura 1) em que a via visual é, à medida que a experiência de leitura aumenta, uma via essencialmente ortográfica.

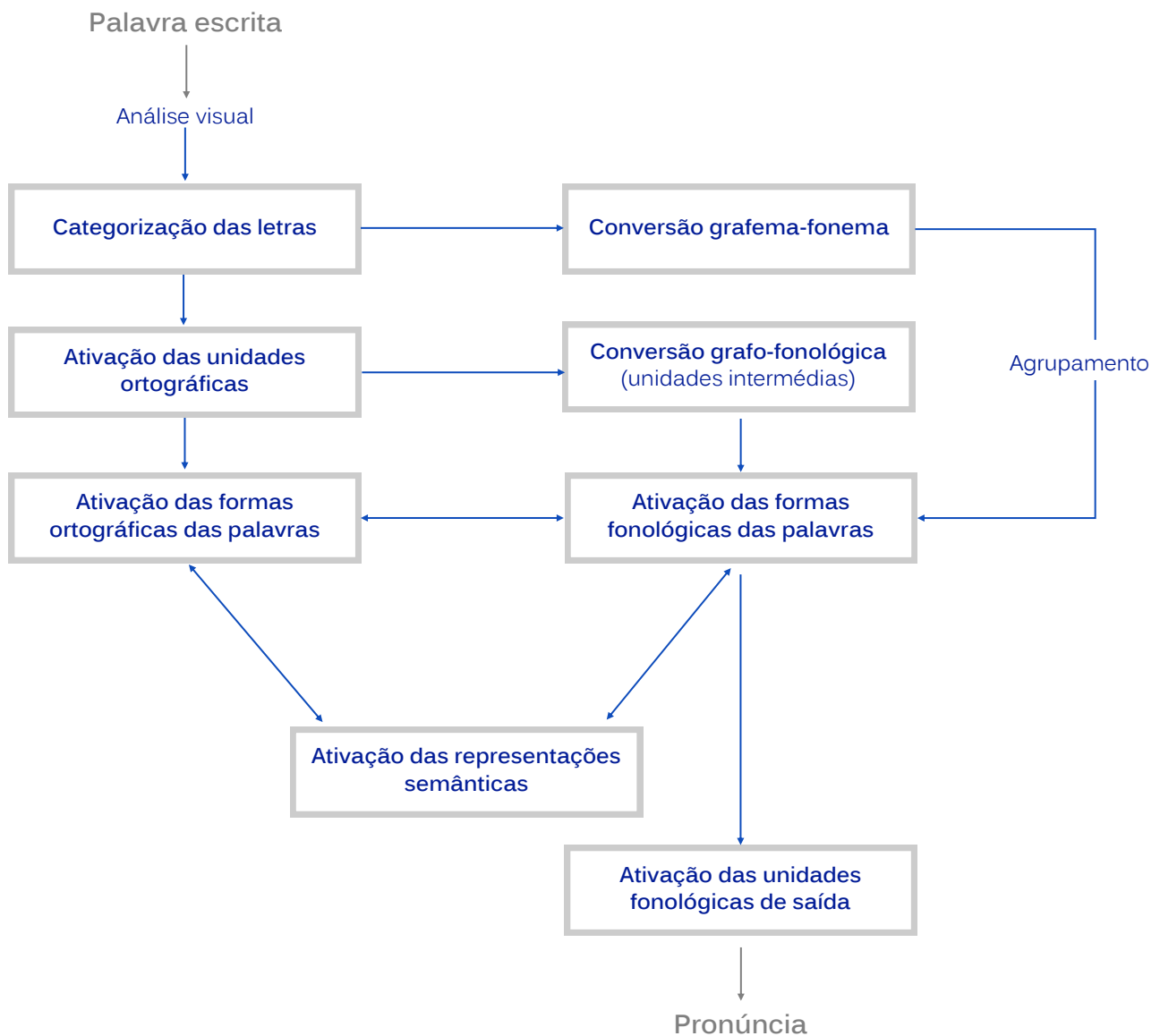


Figura 1. Modelo de Leitura de Palavras (Morais, 1997, p. 128)



Neste modelo, a palavra escrita é sujeita, em primeiro lugar, a uma análise visual, idêntica à de qualquer outro objeto. De seguida, as letras que a constituem são categorizadas (num processo cognitivo complexo) e o resultado desta categorização vai alimentar um sistema de extração de unidades ortográficas e um sistema de conversão grafema-fonema, tornando operacionais estas duas vias.

Na *via da direita* (cf. figura 1), a que o autor chama “fonológica”, as letras são convertidas em sons e, através de um processo de combinação sequencial, ativam as formas fonológicas das palavras que, por sua vez, conduzem à ativação das representações semânticas e ortográficas correspondentes. Ao postular a existência de unidades ortográficas mais amplas do que as letras (ou grafemas), é de admitir que as representações destas unidades ativem diretamente as unidades fonológicas suprafonémicas correspondentes (chamadas unidades intermediárias). O uso de setas de duplo sentido indica que os três tipos de representação das palavras (ortográfica, fonológica e semântica) interagem mutuamente (Morais, 1997).

Na *via da esquerda*, a que o autor chama “ortográfica”, o subsistema que denomina “unidades ortográficas” diz respeito a unidades mais amplas do que a letra (como por exemplo dígrafos, sílabas e partes de palavras como <-era> em *cera* ou em *fera*). São estas unidades ortográficas que ativam as formas ortográficas das palavras e estas, por sua vez, conduzem à ativação das representações semânticas e fonológicas correspondentes. É este processamento, com recurso a unidades ortográficas mais amplas do que a letra, que imprime rapidez à leitura.

Nos primeiros anos de escolaridade o recurso à via fonológica é predominante, pois o aprendiz está ainda a construir o seu léxico ortográfico. À medida que este é construído, a via ortográfica passará a ser cada vez mais usada. O leitor proficiente é aquele que usa a via mais adequada às características das palavras a ler e que serão detalhadas noutro ponto deste capítulo. Os leitores não proficientes e as crianças com dificuldades ao nível da identificação de palavras usam predominantemente uma via, independentemente das palavras a ler.

Todavia, e independentemente do método de ensino da leitura a que é sujeito, o aprendiz descobre muito rapidamente que o sistema de conversão grafema-fonema não é o único de que dispõe para ler e vai integrar pistas semânticas e sintáticas. Aliás, o novo acordo ortográfico (<http://www.priberam.pt/docs/acortog90.pdf>), ao eliminar alguns diacríticos, obriga a atender a pistas sintáticas e semânticas para encontrar a pronúncia correta de algumas palavras (por exemplo, a eliminação do acento circunflexo na palavra *pêlo* ou a eliminação do acento agudo na palavra *pára*).

Fatores que influenciam a leitura de palavras

O léxico mental – conhecimento da língua em que vai aprender a ler

Conhecer (e dominar) a língua em que se aprende a ler é determinante para um acesso bem sucedido à leitura (Carroll, 1987). A revisão da investigação efetuada por McGuinness (2004) evidencia que os maus leitores apresentam défices de linguagem, essencialmente ao nível do vocabulário e do conhecimento morfossintático.

Uma revisão recente da investigação (Hulme & Snowling, 2011) mostra que os leitores que têm dificuldades de compreensão da leitura apresentam défices de linguagem, essencialmente ao nível do vocabulário e do conhecimento morfossintático.

Quer a revisão de estudos efetuada por Anderson e Freebody (1987), quer a realizada por McGuinness (2004) mostraram que as medidas de conhecimento de vocabulário são fortes preditores de uma variedade de índices de competência linguística. A posse de um vocabulário amplo é importante para a leitura, quer ao nível da descodificação, quer ao nível da compreensão. Ao nível da descodificação permite que a criança encontre rapidamente a forma fonológica de uma palavra impressa - por exemplo, a criança pode ler **cá-ssá-có* mas se esta palavra pertencer ao seu vocabulário (léxico mental fonológico e semântico), ela rapidamente emparelha esta forma ortográfica com a forma fonológica conhecida, e lerá corretamente *casaco* (Kɛz'aku), abrindo a vogal <a> da sílaba intermédia e fechando-a na sílaba inicial. Ao nível da compreensão da leitura é talvez mais óbvia a importância do conhecimento do vocabulário. O desconhecimento de uma palavra, nomeadamente se ela for chave de compreensão e não puder ser inferida a partir do contexto, pode comprometer toda a compreensão de um texto.

Possuir um amplo vocabulário não é, todavia, suficiente. É preciso que a criança saiba as combinações possíveis e permitidas pela sua língua, ou seja, é necessário também um bom domínio de competências ao nível morfossintático. As repercussões que problemas a este nível podem ter no processo de leitura não são, no geral, detetadas aquando do início da escolaridade, porque os textos com que as crianças são confrontadas nesta etapa, para além de possuírem um registo muito marcado pela oralidade, apresentam também, no geral, níveis de complexidade muito baixos. Todavia, à medida que a estrutura das frases e dos textos se vai tornando mais complexa, as dificuldades emergem.

Para Tunmer (1990), o domínio morfossintático pode interagir com o desempenho na leitura pelo menos a dois níveis: i) nas possibilidades de o leitor se auto-monitorizar na compreensão do texto; ii) na facilitação da descoberta e apreensão de correspondências grafema-fonema que até aí ignorava. Uma terceira razão poderia, ainda, ser invocada: se o significado de uma ilocução não é um somatório dos significados dos elementos lexicais que a compõem (Gombert, 1992), o acesso ao significado requer uma articulação sintática entre os elementos lexicais isolados. Assim sendo, é evidente que o domínio sintático desempenha um importante papel na extração de sentido.

Conhecimento das correspondências grafema-fonema

A escrita alfabética codifica a fala ao nível do fonema. Isto é, cada fonema de uma palavra é representado na escrita por um grafema (que pode ser simples, como <s> ou <v> ou complexo, como <ch>, <im>, <rr> ou <n>). Por exemplo, a palavra *mar* tem três fonemas – /m/, /a/, /r/ –, por isso a sua escrita obriga à utilização de três grafemas, <m>, <a>, <r>; a palavra *fim* tem dois fonemas – /f/, /ĩ/ –, e, por isso, tem de ser escrita com dois grafemas – <f>, <im>. Para ler com autonomia é indispensável conhecer as correspondências entre grafemas e fonemas, a base do princípio alfabético (Byrne, 1992; Byrne & Fielding-Barnsley, 1989, 1990; Morais,

2012). Sem esse conhecimento não poderão ser lidas as palavras novas que surgem a cada momento, particularmente no início da aprendizagem. Essas correspondências, no entanto, nem sempre são totalmente consistentes, pois há padrões de grafemas, e mesmo grafemas singulares, que podem ser lidos de mais do que uma maneira. O padrão <or> pronuncia-se /ur/ na palavra *por* e /or/ em *cor*; o grafema <c> diz-se /k/ em *cala*, mas pronuncia-se /s/ em *cela*. Estes exemplos mostram que a leitura envolve, para além de conversões grafema – fonema, a aprendizagem de diferentes regras que regem o funcionamento de cada sistema de escrita, os quais são fundamentais para ler palavras com precisão.

Transparência vs opacidade da língua escrita

As línguas de escrita alfabética, mesmo as que usam um mesmo alfabeto, apresentam diferenças em termos da maior ou menor consistência das conversões grafema-fonema. Seymour e colaboradores (2003) apresentam uma proposta de classificação de algumas ortografias europeias, de acordo com esta consistência. Enquanto a ortografia finlandesa é considerada uma ortografia muito regular, a portuguesa aparece classificada num nível intermédio, mais próxima do polo irregular do que a castelhana, a norueguesa, a islandesa, a sueca e a grega. A ortografia inglesa é classificada como a menos regular, antecedida pela dinamarquesa. Esta classificação tem sido sustentada por outras investigações posteriores (Borgwaldt, Hellwig, & de Groot, 2005; Serrano *et al.*, 2010).

Não considerando as letras <k>, <y> e <w>, na língua portuguesa, as 23 letras do alfabeto dão origem a 67 grafemas que representam 35 fonemas (Gomes, 2001). Já no francês, por exemplo, os mesmos 35 fonemas podem ser representados por 130 grafemas (Catach, 1980). No entanto, apesar de a ortografia portuguesa ser considerada de regularidade intermédia em termos de leitura, a descodificação das vogais coloca particulares desafios aos leitores portugueses, dado que as 5 letras vogais correspondem a 14 sons. Por exemplo, o grafema <e> tem uma tradução fonológica irregular, pois pode corresponder a /ɛ/ como em *vela*, /e/ como em *mesa*, /ə/ como em *peru*, ou /i/ como em *errar*. O mesmo acontece com o grafema <x>, já que pode corresponder a /ʃ/ como em *lixo*, /ks/ como em *fixo*, /z/ como em *exame* e /s/ como em *máximo*. Para além desta complexidade, quando o grafema <x> se encontra em fim de sílaba com o grafema <e> transforma-o em ditongo. Numa investigação que analisou a leitura de vogais orais em dissílabos ao longo dos quatro anos do 1.º ciclo (Vale, Sucena, Viana, & Garcez, 2008), verificou-se que vogais menos regulares, <e> e <o>, sofrem um efeito da frequência da palavra a que pertencem. Isto é, quando ocorrem em palavras menos frequentes, como *mega* ou *godo*, dão lugar a níveis de exatidão de cerca de menos 30% a 20%, respetivamente no 1.º e no 4.º anos, do que as mesmas vogais inseridas em palavras frequentes, como *fera* e *lobo*.

A mediação fonológica, isto é, o recurso a um procedimento de transformação de unidades e padrões ortográficos em padrões fonológicos, é necessária para estabelecer e concretizar a pronúncia correta da palavra escrita. Este mecanismo é eficaz para ler uma grande parte das palavras portuguesas, mas não resulta com todas. Por exemplo se essa estratégia fosse aplicada à leitura da palavra *dez*, em vez de /dɛʃ/ obteríamos /deʃ/, já que todas as outras palavras portuguesas terminadas em <-ez> se pronunciam /ej/ no final (Gomes, 2001).

Variáveis psicolinguísticas

Como já foi atrás referido, as palavras não são processadas todas da mesma maneira. Para além das variáveis atrás mencionadas, há ainda um conjunto de outras variáveis relativas às palavras que devem ser consideradas. Entre elas, as mais estudadas são a frequência, a regularidade, a extensão e o tipo de estrutura silábica.

Frequência

A frequência refere-se ao número de vezes que uma palavra ocorre numa língua. Quando uma palavra é lida muitas vezes, o leitor cria uma representação ortográfica consistente dessa palavra, o que permite o seu reconhecimento quase imediato. Estudos realizados com crianças portuguesas mostraram que do 1.º ao 4.º ano de escolaridade as palavras frequentes são lidas com maior precisão e mais depressa do que as palavras infrequentes (Fernandes, Ventura, Querido, & Morais, 2008; Loff & Vale, 2007; Sucena, 2005; Vale, 2000). No início da aprendizagem, o efeito de frequência é tendencialmente maior do que nos níveis de escolaridade mais avançados.

Regularidade /Consistência

A regularidade de uma palavra diz respeito ao grau de precisão com que a pronúncia dessa palavra pode ser derivada das regras de conversão grafema-fonema (RCG-F). Palavras como *fita*, *buzina* ou *pupila* são regulares, pois apenas a última letra pode levantar alguma indecisão quanto à conversão fonológica, embora a posição que ocupa lhe confira uma consistência de 100% relativamente à pronúncia /a/. Por exemplo, os grafemas <f>, <t>, <z>, <p> e <l>, entre outros, não levantam problemas quando aplicadas as regras de conversão grafema-fonema. Já a leitura de palavras como *vacina* e *casulo* implica que seja convocada uma regra contextual (a letra <c> antes de <e> e de <i> lê-se /s/; a letra <s> entre vogais lê-se /z/). Um estudo recente (Sousa, 2011), realizado com crianças portuguesas dos 3.º e 5.º anos de escolaridade, mostrou que a leitura de palavras contendo consoantes cuja pronúncia dependia de uma regra contextual, como *vacina* e *casulo*, dava lugar a mais erros do que a leitura de palavras para as quais podiam ser aplicadas regras estritas de conversão grafema-fonema, como *buzina* e *pupila*. A diferença de desempenho observada entre estes dois tipos de palavras era significativa nos dois anos de escolaridade em análise, embora diminuísse do 3.º para o 5.º ano. Resultados semelhantes tinham também sido verificados por Pinheiro (1995) com crianças brasileiras do 1.º ao 4.º ano de escolaridade.

As palavras *fixo*, *arguido* e *frequente* são irregulares. Na ausência de conhecimentos de ordem etimológica, a sua leitura correta exige o conhecimento da palavra. Para este tipo de palavras, as regras de conversão mais frequentes nem sempre ajudam. Usando os conhecimentos que possui relativamente às palavras que conhece, o leitor pode converter o grafema <x> da palavra *fixo*, erradamente, no fonema /j/, tal como faz para *lixo*. É ainda o conhecimento prévio da palavra que indica ao leitor que o grafema <u> na palavra *frequente* é pronunciado, enquanto não o é na palavra *quente*.



Como foi referido, a identificação, numa língua, de estruturas ortográficas mais amplas do que a letra aumenta a consistência da conversão das vogais e cria regularidades estatísticas que são apreendidas no decurso da aprendizagem (Vale, 2011). Todas as vogais portuguesas são irregulares na leitura (Gomes, 2001), mas o grau de previsibilidade da sua conversão fonológica aumenta drasticamente se tivermos em conta os grafemas adjacentes. Por exemplo, o grafema <e> lê-se /ɛ/ em 75% das vezes que ocorre na estrutura <-eto>, como em *feto*.

Extensão

A extensão das palavras pode ser relativa ao número de letras ou ao número de sílabas que as constituem. Independentemente do critério de extensão usado, verifica-se que as palavras curtas são melhor lidas do que as palavras longas, quando se avalia o tempo de leitura e não apenas a exatidão das respostas (Pinheiro, 1995; Sucena, 2005; Vale, 2000).

Estudos realizados em língua portuguesa que usaram o número de letras como critério de extensão (palavras curtas: 4 a 5 letras; palavras longas: 8 a 9 letras) mostraram que a importância do efeito de comprimento, quer no tempo de leitura, quer na precisão, decrescia à medida que se avançava do 1.º para o 4.º ano de escolaridade. Este padrão de desempenho indica um progressivo afastamento da estratégia de conversão sequencial grafema-fonema e um desenvolvimento gradual dos procedimentos envolvendo conhecimentos lexicais na leitura de palavras (Pinheiro, 1995; Vale, 2000).

Quando a extensão das palavras é medida pelo número de sílabas, o efeito de comprimento parece manifestar-se menos claramente. Num estudo realizado por Sucena (2005), o aumento do número de sílabas não conduziu a um decréscimo na precisão ou no tempo de leitura, quando as palavras incluíam grafemas complexos, como <nh>, <rr>, nem quando continham regras contextuais, como <c> seguido de <e> ou <i>. Já nas palavras constituídas apenas por grafemas simples, os quadrisílabos foram melhor lidos do que os trissílabos e os dissílabos, mas nas palavras irregulares os trissílabos produziram os melhores desempenhos.

A influência da extensão das palavras na leitura também depende da frequência das palavras. Em palavras frequentes a importância do seu comprimento é muito menos relevante. Viana e Ribeiro (2010) não encontraram diferenças ao nível da precisão na leitura de palavras frequentes de duas e três sílabas.

Estrutura silábica

As sílabas são constituídas basicamente por sequências de consoantes (C) e vogais (V). A estrutura silábica CV é a mais frequente, seguida das estruturas CVC, V e CCV que ocupam posições mais ou menos frequentes, dependendo do *corpus* e dos critérios de análise. Estes quatro tipos correspondem a cerca de 85% dos 23 formatos silábicos ortográficos (Gomes, 2001) e a 75% dos 29 tipos de estruturas silábicas faladas (Vigário, Martins, & Frota, 2006).

Existem poucos estudos a documentar a influência da estrutura silábica na qualidade da leitura. Recentemente, um estudo realizado com crianças do 3.º e do 5.º anos de escolaridade (Sousa, 2011) mostrou

que as palavras cuja primeira sílaba tinha uma estrutura CVC, como *castelo* e *silvado*, produziam piores desempenhos do que as palavras com estrutura CV, como *pupila* e *buzina*. Já as palavras com estrutura CCV na primeira sílaba, como *granito* e *flaneta*, deram origem a desempenhos equivalentes aos obtidos com as palavras de estrutura CV, quer no 3.º, quer no 5.º ano. Embora não possamos tirar conclusões a partir de um único estudo, sublinha-se que estes resultados são, em alguns aspetos, semelhantes aos obtidos em estudos realizados com as ortografias francesa e castelhana. Quer no estudo francês (Sprenger-Charolles & Siegel, 1997), quer no estudo espanhol (Defior, Justicia, & Martos, 1996), as palavras com sílabas CV foram melhor lidas do que as palavras com sílabas CVC e CCV. Uma das explicações normalmente apontadas para estas diferenças é a de a estrutura CV ter uma frequência muito superior à das outras estruturas silábicas. No entanto, no caso português, embora a estrutura CCV seja menos frequente do que a CV, os resultados são semelhantes. Provavelmente, fatores como o tipo de vogais que constituem as palavras podem também influenciar os resultados, pelo que serão necessários estudos que verifiquem esta hipótese.

A descodificação: Implicações para a sua avaliação

A descodificação é o mecanismo responsável pela conversão de cada grafema de uma dada palavra num fonema, que fica ativo na memória de trabalho e que, combinado com os outros, vai permitir formar a pronúncia de uma palavra. O facto de grande parte das dificuldades na aprendizagem da leitura ocorrer ao nível do tratamento de palavras traduz a importância deste processo e justifica a sua avaliação sistemática e precoce.

Como foi referido no capítulo anterior, os procedimentos para avaliar a leitura de palavras podem ser integrados em dois grandes grupos: a) o recurso a palavras apresentadas de forma isolada; b) a leitura de palavras em (con)texto.

O paradigma experimental da leitura de “palavras isoladas”, isto é, de palavras que não estão inseridas num contexto sintático e semântico, tem sido utilizado por três razões fundamentais:

- 1) Porque sendo apresentadas isoladamente não sofrem a influência das palavras adjacentes. Quando tem de ler palavras isoladas, isto é, fora de um contexto, o leitor não pode gerar expectativas sobre qual a palavra que está escrita, por isso tem de usar os conhecimentos que possui sobre o funcionamento do sistema alfabético para produzir a pronúncia correta, a partir do padrão ortográfico que lhe é apresentado.
- 2) Porque a leitura de palavras isoladas é o procedimento privilegiado para estudar o tipo de estratégias e mecanismos usados na leitura (Peereman, 1992) e é, por isso, considerada uma medida mais rigorosa das capacidades de leitura do que a leitura em (con)texto.
- 3) Porque a leitura de palavras isoladas tem revelado ser o preditor mais robusto do sucesso ulterior em leitura (Stuart, 1995; Verhoeven & van Leeuwe, 2009). Para além de as diferenças individuais em leitura de palavras serem, logo desde o 1.º ano, fortemente preditivas dos desempenhos mais tardios na leitura, o nível de

desempenho nessa tarefa é também uma medida que apresenta grande estabilidade individual ao longo de vários anos de escolaridade (Foorman, Francis, Shaywitz, Shaywitz, & Fletcher, 1997; Verhoeven & van Leeuwe, 2009).

A construção do TLP (Teste de Leitura de Palavras)

No capítulo I foram abordados os dois procedimentos básicos de avaliação da descodificação: a leitura de palavras apresentadas de forma isolada e a leitura de palavras apresentadas em (con)texto. Neste capítulo, foram analisadas as principais razões para o uso da leitura de palavras apresentadas de forma isolada.

Dados os objetivos da BAL, o teste de leitura de palavras que a integra – o TLP – responde de forma simples, e facilmente aplicável em contextos educativos, à necessidade de avaliar o desempenho dos alunos ao nível da leitura de palavras. A necessidade desta avaliação justifica-se por se tratar de uma área na qual um número significativo de alunos apresenta dificuldades e para os quais é necessário efetuar um diagnóstico precoce, tendo em vista a organização de um plano de intervenção. Uma segunda razão prende-se com a relação entre descodificação e as dificuldades de compreensão da leitura. Como foi referido no capítulo I, e será retomado no capítulo III, quando estamos em presença de alunos com dificuldades de compreensão da leitura, é necessário determinar em que medida as mesmas são uma consequência de problemas ao nível da descodificação. Responder a esta questão é crítico para definir objetivos de intervenção.

Na BAL optou-se por incluir uma prova de leitura de palavras apresentadas de forma isolada em formato digital, permitindo a obtenção de informação sobre a precisão da leitura de cada palavra. Os critérios que presidiram à seleção do banco de palavras, e que serão explicitados no capítulo IV, foram os mais usados na investigação relativos às características psicolinguísticas das palavras: frequência, regularidade, extensão e estrutura silábica.

Referências bibliográficas

- Anderson, R. C. & Freebody, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE; International Reading Association.
- Borgwaldt, S. R., Hellwig, F. M., & de Groot, A. M. B. (2005). Onset entropy matters: Letter-to-phoneme mappings in seven languages. *Reading and Writing, 18*, 211-229.
- Byrne, B. (1992) Studies in the acquisition procedure for reading: Rationale, hypotheses, and data. In P. B. Gough, L. C. Ehri & R. Treiman (Eds.), *Reading acquisition* (pp. 1-34). Hillsdale, NJ, USA: Lawrence Erlbaum Associates.
- Byrne, B. & Fielding-Barnsley, R. (1989). Phonemic awareness and letter knowledge in the child's acquisitions of the alphabetic principle. *Journal of Educational Psychology, 80*(1), 313-321.

- Byrne, B. & Fielding-Barnsley, R. (1990). Acquiring the alphabetic principle: A case For teaching recognition of phoneme identity. *Journal of Educational Psychology*, 82, 805-812.
- Carroll, J. B. (1987). The nature of the reading process. In H. Singer & R. B. Rudell (Eds.), *Theoretical models and processes of reading* (3rd Ed.) (pp. 292-303). Newark, DE: International Reading Association.
- Catach, N. (1980). *L'orthographe française*. Paris: Nathan.
- Defior, S., Justicia, F., & Martos, F. (1996). The influence of lexical and sublexical variables in normal and poor Spanish readers. *Reading and Writing: An Interdisciplinary Journal*, 8(6), 487-497.
- Ehri, L. C. (2005). Learning to read words: Theory, findings and issues. *Scientific Studies of Reading*, 9, 167-189.
- Ellis, A. W. (1989). *Lecture, écriture et dyslexie: Une approche cognitive*. Neuchâtel, Paris: Delachaux & Niestlé.
- Fernandes, S., Ventura, P., Querido, L., & Morais, J. (2008). Reading and spelling acquisition in European Portuguese: A preliminary study. *Reading and Writing: An Interdisciplinary Journal*, 21, 805-821
- Foorman, B. R., Francis, D. J., Shaywitz, S. E., Shaywitz, B. A., & Fletcher, J. M. (1997). The case for early reading intervention. In B. A. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 243- 264). Mahwah, NJ, USA: Lawrence Erlbaum Associates.
- Gombert, J. E. (1992). Activité de lecture et activités associées. In M. Fayol, J. E. Gombert, P. Lecocq, L. Sprenger-Charolles & Daniel Zagar (Eds.), *Psychologie cognitive de la lecture* (pp. 107-140). Paris: P.U.F.
- Gomes, I. (2001). *Ler e escrever em português europeu*. Tese de Doutoramento, não publicada. Porto: Faculdade de Psicologia e Ciências da Educação da Universidade do Porto.
- Goodman, K. S. (1968). *The psycholinguistic nature of the reading process*. Detroit: Wayne State University Press.
- Gough, P. B. (1972). One second of reading. In J. F. Kavanaugh & I. G. Mattingly (Eds.), *The relationship between speech and hearing* (pp. 331-358). Cambridge, MA: MIT Press.
- Høien-Tengesdal, I. & Tønnessen, F-E. (2011). The relationship between phonological skills and word decoding. *Scandinavian Journal of Psychology*, 52, 93-103. doi: 10.1111/j.1467-9450.2010.00856.x.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293-323.
- Loff, A. & Vale, A. P. (2007). A análise do padrão de erros de leitura de palavras em língua portuguesa nos quatro anos do 1.º Ciclo do Ensino Básico. In A. Barca, M. Peralbo, A. Porto, B. D. Silva & L. Almeida (Eds.), *Revista Galego-Portuguesa de Psicología e Educación, Número Extraordinário*, 1047-1059.
- McGuinness, D. (2004). *Language development and learning to read: The scientific study of how language development affects reading skill*. Cambridge, Mass.: MIT Press.

- McGuinness, D. (2006). *O ensino da leitura. O que a ciência nos diz sobre como ensinar a ler*. Porto Alegre: Artmed.
- Morais, J. (1997). *A arte de ler. Psicologia cognitiva da leitura*. Lisboa: Cosmos.
- Morais, J. (2012). *Criar leitores. O ensino da leitura – para professores e encarregados de educação*. Porto: Livpsic.
- Peereman, R. (1992). Lecture, écriture, orthographes. In P. Lecocq (Ed.), *La lecture: Processus, apprentissage, troubles* (pp. 55-78). Lille: Presses Universitaire de Lille.
- Pinheiro, A. (1995). Reading and spelling development in Brazilian Portuguese. *Reading & Writing*, 7(1), 111-138.
- Rumelhart, D. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI* (pp. 573-603). Hillsdale, NJ: Erlbaum.
- Rumelhart, D. E. (1994). Toward an interactive model of reading. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed.), (pp. 864-894). Newark, DE: International Reading Association.
- Samuels, S. J. (1994). Toward a Theory of Automatic Information Processing in Reading, Revisited. In R. B. Ruddell, M.R. Ruddell, and H. Singer (Eds.). *Theoretical models and processes of reading* (pp. 816-837). Newark, DE: International Reading Association.
- Serrano, F., Genard, N., Sucena, A., Defior, S., Alegria, J., Mousty, P., (...), & Seymour, Ph. (2010). Variations in reading and spelling acquisition in Portuguese, French and Spanish: A cross-linguistic comparison. In S. L. Castro, & J. Morais (Eds.), Special issue: Lexical processing: Phonetic, prosodic and morphological aspects. *Journal of Portuguese Linguistics*, 9-10, 183-205.
- Seymour, P. H. K., Aro, M., Erskine, J. M., Wimmer, H., Leybaert, J., Elbro, C., (...), & Olofsson, A. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143-174.
- Smith, F. (1973). *Psycholinguistics and reading*. New York: Holt, Rinehart & Winston.
- Sousa, J. (2011). *Dislexia: Atraso ou desvio no desenvolvimento?* Dissertação de Mestrado, não publicada. Vila Real: Universidade de Trás-os-Montes e Alto Douro.
- Sprenger-Charolles, L., & Siegel, L. (1997). A longitudinal study of the effects of syllabic structure on the development of reading and spelling skills in French. *Applied Psycholinguistics*, 18(4), 485-505. doi: <http://dx.doi.org/10.1017/S014271640001095X>
- Snowling, M. & Hulme, C. (2005) (Eds.). *The science of reading: A handbook*. Oxford: Blackwell Publishing.
- Stanovich, K. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16, 32-71.
- Stuart, M. (1995). Prediction and qualitative assessment of five- and six-year-old children's reading: A longitudinal study. *British Journal of Educational Psychology*, 65, 287-296.

- Sucena, A. (2005). *Aprendizagem da leitura e da escrita em português Europeu numa perspetiva translinguística*. Tese de Doutoramento, não publicada. Porto: Faculdade de Psicologia e Ciências da Educação da Universidade do Porto.
- Tunmer, W. E. (1990). The role of language prediction skills in beginning reading. *New Zealand Journal of Educational Studies*, 25, 95-114.
- Vale, A. P. (2000). *Correlatos metafonológicos e estratégias iniciais de leitura-escrita de palavras no português: Uma contribuição experimental*. Tese de Doutoramento, não publicada. Vila Real: Universidade de Trás-os-Montes e Alto Douro.
- Vale, A. P. (2011). Orthographic context-sensitivity in vowel decoding by Portuguese monolingual and Portuguese-English bilingual children. *Journal of Research in Reading*, 34(1), 43-58. doi: 10.1111/j.1467-9817.2010.01482.x. ISSN: 0141-0423.
- Vale, A. P., Sucena, A., Viana, F., & Garcez, A. (2008). Desenvolvimento do processo de descodificação de grafemas vocálicos em português europeu. Comunicação apresentada no 3.º Encontro Nacional da Associação Portuguesa de Psicologia Experimental. Portugal: Universidade do Algarve, Faro, 28 e 29 de março 2008.
- Verhoeven, L. & van Leeuwe, J. (2009). Modeling the growth of word-decoding skills: Evidence from Dutch. *Scientific Studies of Reading*, 13(3), 205-223. doi: 10.1080/10888430902851356.
- Viana, F. L. & Ribeiro, I. S. (2010). *PRP - Prova de Reconhecimento de Palavras*. Lisboa: CEGOC-TEA.
- Vigário, M., Martins, F., & Frota, S. (2006). A ferramenta FreP e a frequência de tipos silábicos e classes de segmentos no Português. In F. Oliveira & J. Barbosa (Eds.), *XXI Encontro da Associação Portuguesa de Linguística. Textos Seleccionados* (pp. 675-687). Lisboa: Associação Portuguesa de Linguística. Acedido em 20 de junho 2013: <http://ww3.fl.ul.pt/LaboratorioFonetica/frep/bef/APL2006VigarioMartinsFrota.pdf>

COMPREENSÃO DE TEXTOS. CONSIDERAÇÕES TEÓRICAS E IMPLICAÇÕES PARA A SUA AVALIAÇÃO

Alina Galvão Spinillo, Iolanda Ribeiro e Fernanda Leopoldina Viana

O conceito de compreensão aparece frequentemente associado à leitura, porém não se esgota nesta. Aplica-se sempre que o sujeito tem de processar um texto, sendo que este pode ser apresentado, quer numa modalidade oral, quer numa modalidade impressa. Em ambos os casos a compreensão é crítica para os sujeitos em todos os momentos da sua vida, uma vez que ela remete para a capacidade de processar a informação proveniente do mundo. Em contexto escolar esta competência tem um valor particular, uma vez que, sem serem capazes de extrair e construir significados, os alunos não conseguem apropriar-se das aprendizagens associadas às várias áreas curriculares.

As questões da avaliação da compreensão colocam-se, por isso, com uma especial relevância. Para os profissionais de diferentes áreas que lidam com alunos que têm problemas de aprendizagem, sejam estes na esfera mais restrita da leitura ou noutras áreas académicas, é absolutamente necessário poderem dispor de testes que contribuam para um diagnóstico preciso dos problemas de compreensão.

Reconhecer a relevância da compreensão e da sua avaliação é mais simples do que lidar com duas questões fundamentais no que concerne à sua medida: a) Qual o enfoque teórico a adotar? e b) Como avaliar? A primeira questão conduz à necessidade de definir e delimitar o conceito, enquanto a segunda orienta para a necessidade de tomar decisões sobre as tarefas a usar, o modo de as apresentar (leitura *versus* oral) e a escolha de textos (narrativos *versus* informativos). A resposta às questões formuladas justifica as opções tomadas na construção das provas de compreensão da BAL – Bateria de Avaliação de Leitura

Compreensão de textos: considerações teóricas

Compreender textos, sejam orais ou escritos, requer considerar o texto e o indivíduo que com ele interage em contextos sociais diversos. O texto pode ser apresentado oralmente ou por escrito, havendo aspetos comuns à compreensão oral e escrita e aspetos específicos particulares a uma ou outra modalidade de apresentação do texto. Por exemplo, a leitura de textos envolve habilidade de decodificação, habilidade esta que não é requerida na compreensão de textos apresentados oralmente. Para compreender um texto, o leitor aciona um conjunto de habilidades cognitivas (memória, capacidade de monitoramento, capacidade de estabelecer inferências) e linguísticas (habilidade de decodificação – no caso da leitura de textos escritos, vocabulário, conhecimentos sobre a sintaxe e sobre os géneros textuais, entre outros) e convoca o seu conhecimento do mundo (informações acumuladas relacionadas com as experiências sociais vivenciadas, variando desde conhecimentos gerais até conhecimentos específicos adquiridos formal ou informalmente). Na realidade, a



compreensão de textos é um fenômeno multifacetado que envolve três dimensões distintas, porém indissociáveis: a linguística, a cognitiva e a social (Spinillo, 2013). Estas dimensões, que estão sempre presentes na compreensão, assumem um papel de maior ou menor destaque em função da perspectiva teórica adotada por quem se propõe investigar este tema. Além disso, a cada uma destas dimensões estão associados fatores que são responsáveis e que influenciam a compreensão de textos, como será discutido adiante, após a apresentação dos diferentes modelos teóricos que procuram definir e descrever a compreensão textual.

Compreensão de textos e diferentes modelos teóricos

Definir a compreensão de textos não é tarefa fácil devido à sua complexidade e natureza multifacetada. Para tal, é necessário considerar os modelos teóricos que procuram explicar como a compreensão ocorre e considerar os fatores que a influenciam. De forma breve, como referem Colomer e Camps (2002), há três grandes classes de modelos relativos ao processamento de leitura de um texto: modelos de processamento ascendente, modelos de processamento descendente e modelos interativos.

Os *modelos de processamento ascendente* pressupõem que a compreensão se inicia a partir de níveis inferiores do texto (decodificação, reconhecimento de palavras) até alcançar os níveis superiores (orações, frases, parágrafos, texto como um todo). De acordo com este modelo, inicialmente, o leitor seleciona, de entre os vários sentidos que uma palavra pode ter, aquele que melhor se adequa ao texto. O leitor constrói, então, uma rede de relações semânticas e sintáticas ao nível das orações ou frases, extraindo o significado a partir da sua estrutura gramatical. As frases são, então, relacionadas entre si, formando uma representação mental de todo o texto. Assim que um dado segmento do texto é processado, ele é integrado no que foi anteriormente apresentado (e que se encontra armazenado na memória de trabalho), de modo que cada nova informação é integrada nas anteriores. Esta integração é responsável pelos significados que serão atribuídos ao próximo segmento de texto. Neste modelo, são fortemente enfatizadas as relações intratextuais estabelecidas durante a leitura.

Os *modelos de processamento descendente* pressupõem que a compreensão se inicia com o contexto em que o leitor interage com o texto: a situação na qual a leitura ocorre, os propósitos da leitura, as suas expectativas diante do texto e o seu conhecimento do mundo. É a partir deste cenário que o indivíduo vai aceder ao texto e atribuir significados a todos os seus segmentos (palavra, frase, parágrafo, ...). Neste modelo, a ênfase recai sobre relações extratextuais estabelecidas durante a leitura.

Os *modelos interativos* pressupõem que a compreensão envolve tanto um processamento ascendente como descendente em todos os níveis do texto (palavra, frase, parágrafo), de modo que o conhecimento do mundo e o conhecimento linguístico do leitor participam e interagem simultaneamente em todos os níveis de processamento do texto.

Um dos modelos mais aceites entre os investigadores é o modelo denominado Construção-Integração (Kintsch, 1998; Kintsch & Rawson, 2005) que se define como sendo interativo. A arquitetura deste modelo baseia-se em duas instâncias que são o texto-base e o modelo situacional. O texto-base consiste numa

representação mental muito próxima do texto lido, marcada por informações intratextuais; o modelo situacional é uma representação mental marcada pelo conhecimento do mundo que o leitor possui. O modelo pressupõe ainda fases cíclicas de construção e de integração, nas quais, a partir de conexões entre as passagens do texto, o leitor constrói um modelo mental a nível local (atribuindo significado às palavras e às proposições), mas também a nível mais global, ao integrar nas informações intratextuais o seu conhecimento de mundo, gerando novas informações, que são as inferências. Assim, representação mental e inferências surgem como instâncias responsáveis por um processamento de natureza ascendente e de natureza descendente no momento da leitura.

Tomando por base uma perspectiva interativa, a compreensão de textos pode ser definida como sendo um processo inferencial por excelência, no qual o leitor tem um papel ativo na interpretação do texto, uma vez que este processo não se limita apenas à atividade de identificação ou de reprodução da informação, sendo um processo que implica a construção e a apropriação de sentidos, tendo como base a relação entre o autor, o texto e o próprio leitor. Para construir sentidos, o leitor interpreta progressivamente o texto, gerando hipóteses à medida que lê, as quais serão confirmadas, infirmadas ou modificadas, ao longo da leitura. Percebe-se, assim, a natureza dinâmica deste processo, o qual requer uma constante elaboração e reelaboração por parte do leitor (Solé, 1998).

Ressalta-se aqui, que a construção de uma representação mental e o estabelecimento de inferências são o cerne da compreensão de textos, constituindo os processos responsáveis, tanto por um processamento de natureza ascendente como pelo de natureza descendente no momento da leitura.

Os fatores que influenciam a compreensão de textos

Como mencionado, e aceite por um número expressivo de estudiosos da área, compreender textos envolve fatores sociais, linguísticos e cognitivos. Os fatores sociais, de maneira geral, estão relacionados, tanto com o contexto em que a interação entre o leitor e o texto ocorre (objetivos, motivações e expectativas do leitor e do autor do texto), como com o contexto sociocultural do leitor, as suas experiências vividas e acumuladas (Dell'Isola, 2001; Kleiman, 2002a, 2002b; Koch & Elias, 2006; Marcuschi, 2008; Solé, 1998; Tolchinsky & Pipkin, 2003).

Os fatores linguísticos referem-se ao domínio que o leitor possui acerca dos aspetos lexicais, sintáticos e semânticos, bem como à habilidade de descodificação (Cain & Oakhill, 2004; Kleiman, 2002a; 2002b). Por exemplo, dificuldades acentuadas na descodificação podem comprometer o reconhecimento da palavra, a fluência e o estabelecimento de relações entre os diversos segmentos textuais (palavras, frases, parágrafos), prejudicando o processo inferencial e a construção de uma representação global do texto (Balota, 1990; Ehri, 1992; Salles & Parente, 2002, 2004; Yuill & Oakhill, 1991). O vocabulário, por sua vez, é um dos preditores da compreensão, como mostram diversos estudos (Perfetti, 1985; Stanovich, West, Cunningham, Cipielewski, & Siddiqui, 1996). Na realidade, o vocabulário está relacionado com o conhecimento do mundo, de modo que um vocabulário limitado dificulta a atribuição de significados apropriados ao texto.

Para além do processamento semântico, também o processamento sintático desempenha um papel importante, visto que o conhecimento que o leitor possui acerca das categorias sintáticas e das regras gramaticais da língua é acionado para atribuir significado às orações e frases. Uma vez compreendidas, estas são integradas noutras orações e frases, gerando uma interpretação local e global do texto. Conhecimentos sobre a estrutura do texto e sobre os seus elementos constituintes também participam no complexo processo de compreensão, podendo comprometê-lo (Solé, 2003). Correlações entre o conhecimento sobre a organização geral do texto e o nível de compreensão foram documentadas por Cain e Oakhill (1996; 2004). De acordo com as autoras, quanto maior for o domínio acerca da organização e constituição dos textos mais condições o leitor terá para os compreender. Isto ocorre porque a familiaridade com a estrutura do texto auxilia na integração de informações, facilitando a construção de um modelo mental e o estabelecimento de inferências.

Os fatores cognitivos referem-se à memória (de trabalho e de longo prazo), à monitorização da compreensão e ao estabelecimento de inferências, sendo referidos na literatura como processos de alto nível (Oakhill & Yuill, 1996). A memória de trabalho atua na retenção de informações intratextuais durante a leitura, permitindo que o leitor integre as diferentes passagens do texto e, assim, estabeleça inferências, monitorize a sua compreensão e construa um modelo mental de todo o texto, tanto a nível local como a nível global (Fletcher, van der Broek, & Arthur, 1996; Kintsch, 1998; Oakhill & Yuill, 1996; Yuill & Oakhill, 1991). A memória de longo prazo permite que o leitor acione tanto o seu conhecimento do mundo como o seu conhecimento linguístico. A monitorização consiste na tomada de consciência a respeito de sua própria compreensão, permitindo que o leitor identifique o que impede ou dificulta a compreensão e que identifique quando e onde, eventualmente, deixou de compreender. Os estudos conduzidos por Perfetti, Marron e Foltz (1996) e por Ruffman (1996) mostram que o bom leitor é capaz de monitorizar a sua compreensão e de utilizar estratégias de leitura que o ajudam a superar as dificuldades identificadas, como, por exemplo, fazer correções, ler devagar, ler novamente uma dada passagem, procurar outras fontes de informação etc. Nada disso é realizado por leitores menos hábeis, que, com frequência, nem sequer se apercebem das dificuldades que experimentam. A importância da monitorização para a compreensão de textos foi enfatizada por Coelho e Correa (2010), a partir de um estudo de intervenção junto de adolescentes que foram estimulados a monitorizar a sua leitura e a identificar inconsistências em textos. Estes adolescentes tiveram um desempenho significativamente superior na compreensão de textos, quando comparados com um grupo equivalente de adolescentes que não tinham usufruído desta experiência.

O estabelecimento de inferências é crucial para a compreensão, pois é através delas que o leitor preenche as lacunas do texto com o conhecimento do mundo (Brandão & Oakhill, 2005; Colomer & Camps, 2002; Coscarelli, 2003; Graesser & Zwaan, 1995; Kintsch, 1998; Marcuschi, 2008; Oakhill & Yuill, 1996; Spinillo & Mahon, 2007; Vidal-Abarca & Rico, 2003; Yuill & Oakhill, 1991). As inferências vão além daquilo que está explicitado no texto, permitindo gerar novas proposições a partir das informações dadas. São elas que possibilitam a integração de informações intratextuais entre si e de informações intratextuais com informações extratextuais. Leitores com dificuldades de compreensão tendem a memorizar as informações literais em detrimento das inferenciais,

como demonstrado por Yuill e Oakhill (1991). Autores de diferentes afiliações teóricas afirmam que o leitor proficiente é aquele que é capaz de fazer uma leitura profunda e ir além do que está explicitado no texto. Os níveis de compreensão textual serão apresentados e discutidos adiante.

Segundo Spinillo (2013), a compreensão de textos pode também ser influenciada por fatores circunstanciais, uma vez que leitores com diferentes propósitos podem diferir em relação à compreensão de um mesmo texto, ou um mesmo leitor pode ter diferentes níveis de compreensão do mesmo texto, ao lê-lo em diferentes momentos. Os propósitos do leitor também podem variar em função do tipo de texto. Por exemplo, ao ser informado de que lerá um conto de fadas, o leitor cria a expectativa de que o conteúdo do texto é fictício; por outro lado, ao ser informado de que lerá uma notícia de jornal, ele tem a expectativa de que o conteúdo seja real. Além disso, os aspetos internos do texto direcionam a compreensão, antecipando o que se segue. Por exemplo, numa história, o título pode remeter para a ação ou para as personagens do texto, enquanto o título de uma secção do jornal pode antecipar o tipo de conteúdo que será tratado (política, economia, desporto, cinema etc.). De acordo com Solé (1998; 2003), ao saber qual o tipo de texto que será lido e ao estar ciente dos propósitos da leitura, o leitor cria um esquema mental que favorece a integração de informações e o estabelecimento de inferências.

Os níveis de compreensão de textos

Tomando por base a capacidade de estabelecer inferências, diversos autores têm procurado descrever níveis de compreensão de textos. Marcuschi (2008), por exemplo, especifica horizontes de compreensão que caracterizam a relação do leitor com o texto: falta de horizonte, horizonte mínimo, horizonte máximo, horizonte problemático e horizonte indevido. A falta de horizonte caracteriza-se pela repetição das informações literais presentes no texto e o horizonte mínimo consiste numa paráfrase das informações. Assim, ambos os horizontes consistem em níveis de compreensão que se restringem a informações literais presentes no texto. O horizonte máximo caracteriza-se pelo estabelecimento de inferências apropriadas, indo além do que é explicitamente mencionado no texto. O horizonte problemático vai muito além do que é explicitado no texto, extrapolando-o, pois o leitor insere muitos elementos de conhecimento pessoal e de opinião. O horizonte indevido afasta-se dos sentidos autorizados pelo texto.

King (2007) propõe, por sua vez, a distinção entre compreensão literal ou superficial e compreensão inferencial ou profunda. A primeira baseia-se nas informações que estão explícitas no texto, de modo que a origem das inferências é o texto, havendo pouca elaboração por parte do leitor. A segunda – a compreensão profunda – baseia-se em conhecimentos que vão além do texto e que derivam do estabelecimento de inferências que dependem do papel ativo desempenhado pelo leitor ao acionar conhecimentos do mundo que são pertinentes para o entendimento do conteúdo do texto.

Compreensão de textos: implicações para a sua avaliação

A compreensão de textos depende, como referimos previamente, de uma constelação de competências cognitivas, linguísticas e sociais. Algumas são específicas à compreensão da leitura (como a descodificação), outras, como a memória de trabalho ou o conhecimento lexical, são variáveis que explicam a compreensão de textos, qualquer que seja a modalidade de apresentação (leitura ou oral), mas também outros comportamentos e processos psicológicos. A avaliação destas variáveis é passível de ser efetuada a partir de um conjunto diverso de provas psicológicas. A título ilustrativo, a memória de trabalho pode ser avaliada recorrendo ao teste de memória de dígitos da WISC – Escala de Inteligência de Wechsler para Crianças (Wechsler, 2003). Por sua vez, a prova de vocabulário da WISC ou as provas de nomeação e de definição verbal da bateria de provas de Avaliação da Linguagem Oral (Sim-Sim, 2001) permitem a avaliação do vocabulário. Não se pretende, neste capítulo, fazer uma sistematização das várias abordagens e procedimentos que podem ser usados no quadro da avaliação dos fatores que explicam as diferenças individuais ao nível da compreensão e que deverão ser especificamente contemplados, quando se trata, por exemplo, de avaliar crianças sinalizadas por apresentarem problemas de compreensão.

Os modelos revistos ao longo deste capítulo foram conceptualizados, tendo por base a leitura de textos escritos. Todavia, e como foi referido no capítulo I, os modelos e teorias relativos à compreensão da leitura não estabelecem uma distinção clara entre compreensão da linguagem oral e compreensão da linguagem escrita (Cain & Oakhill, 2008), considerando que o produto é idêntico – uma representação mental coerente, integrada e completa (Kintsch, 1998). Ou seja, os processos para construir as representações mentais são os mesmos, independentemente da modalidade de apresentação ser escrita ou falada (Gernsbacher, 1990). Na BAL optou-se, pelas razões discutidas no capítulo I, por avaliar a compreensão, apresentando ao sujeito os textos através de duas modalidades: oral e escrita.

O conhecimento dos modelos teóricos de compreensão é fundamental para os profissionais, psicólogos, professores ou outros especialistas que trabalham com crianças que apresentam dificuldades em atribuir significado ao que leem ou ouvem ler, uma vez que facultam um quadro de referência compreensivo que contribui para a formulação de hipóteses sobre a fonte das dificuldades. Do ponto de vista da medida, os modelos teóricos são fundamentais porque orientam a decisão do investigador sobre o “que” e “como” avaliar. Na BAL optou-se por avaliar níveis de compreensão.

São várias as tentativas de classificação dos diferentes níveis de compreensão da leitura¹. Destas tentativas (Barrett, 1976; Català, *et al.*, 2001; Herber, 1978; Pearson & Johnson, 1978) resultaram diferentes propostas, nas quais os autores optam por terminologias diferentes (compreensão literal, compreensão inferencial, reorganização, compreensão crítica, compreensão avaliativa, compreensão interpretativa, compreensão aplicada, entre outros) e com um número variável de níveis (entre 3 e 5). Acresce ainda que alguns autores se

¹ Para uma revisão detalhada consultar Cadime, 2012.

centraram unicamente nas inferências, construindo taxonomias para as classificar (e.g. Cunningham, 1987; Graesser, Singer, & Trabasso, 1994; Graesser, Swamer, Baggett, & Sell, 1996; Giasson, 2005). Embora ao longo deste capítulo se tenha enfatizado a importância das inferências, tal não significa que a localização de informação explícita presente num texto não seja igualmente importante.

Esta breve revisão mostra que a classificação dos níveis de compreensão está longe de ser consensual, pelo que se impõe ao investigador que quer construir provas de avaliação da compreensão de textos tomar decisões quanto a este aspeto. Na BAL optou-se por definir quatro níveis de compreensão, baseados na releitura das propostas de Barrett (1976) efetuada por Català e colaboradores (2001), dado que esta nomenclatura é a mais próxima da utilizada nos programas de Língua Portuguesa, nos Estudos do PISA e nas provas de aferição de Língua Portuguesa. A compreensão literal implica a identificação de ideias e de informação explícita. A compreensão inferencial resulta da integração da informação veiculada pelo texto e das experiências e conhecimentos do leitor, conduzindo à formulação de hipóteses ou inferências. A reorganização está relacionada com tarefas de análise, de síntese e/ou de organização da informação presentes no texto de forma a aceder a conhecimento novo. Por fim, a compreensão crítica requer a formulação de juízos por comparação com um critério. Estes juízos podem incluir, por exemplo, distinções entre realidade e fantasia, factos e opiniões.

As diferentes taxonomias de compreensão da leitura não consideram distinções em função do tipo de texto. Todavia, na BAL, e pelas razões discutidas no capítulo I, apresentam-se testes separados em função do tipo de texto usado.



A construção das provas de compreensão: Testes de Compreensão de Textos (narrativos e informativos) na modalidade de leitura e Testes de Compreensão de Textos (narrativos e informativos) na modalidade oral

No capítulo I descreveram-se as principais dificuldades na leitura que os alunos podem apresentar. Existe uma discussão em aberto sobre a classificação daquelas dificuldades, bem como sobre as terminologias a usar. Não é propósito deste livro debruçar-se sobre esta temática. A análise é efetuada a partir das áreas nas quais os alunos apresentam dificuldades, podendo ser descritos os seguintes perfis:

- a) Alunos com dificuldades na leitura de palavras, na compreensão da linguagem oral e na compreensão da leitura;
- b) Alunos sem dificuldades na leitura de palavras, mas com dificuldades na compreensão da linguagem oral e na compreensão da leitura;
- c) Alunos com dificuldades na leitura de palavras e sem dificuldades na compreensão da linguagem oral. Este grupo pode apresentar dificuldades na compreensão da leitura que decorrem de problemas de descodificação.

A caracterização de cada um destes perfis requer que os alunos sejam avaliados através de provas que avaliem a leitura de palavras, a compreensão da linguagem oral e a compreensão da leitura. Têm sido descritos perfis mais complexos, que contemplam outras variáveis (e.g. vocabulário, fluência da leitura) (Defior, 2000; Nation, 2005; Spear-Swerling & Sternberg, 1996). No entanto, os perfis aqui descritos são particularmente úteis para um diagnóstico das áreas de dificuldade com implicações diretas para a definição de objetivos de intervenção e de organização de respostas adequadas às necessidades dos alunos. Como referido ao longo deste capítulo, as diferenças de desempenho na compreensão manifestam-se em função dos textos usados. A integração desta informação justifica a construção de quatro testes independentes que avaliam a compreensão de textos narrativos e informativos apresentados em duas modalidades que requerem a audição e a leitura:

- ◆ Teste de compreensão de textos na modalidade oral – narrativo (TCTMO-n);
- ◆ Teste de compreensão de textos na modalidade oral – informativo (TCTMO-i);
- ◆ Teste de compreensão de textos na modalidade de leitura – narrativo (TCTML-n);
- ◆ Teste de compreensão de textos na modalidade de leitura – informativo (TCTML-i).

Referências bibliográficas

Balota, D. (1990). The role of meaning in word recognition. /n D. Balota, G. Flores D' Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9-32). Hillsdale, NJ.: Lawrence Erlbaum Associates.

- Barrett, T. C. (1976). Taxonomy of reading comprehension. In R. S. T. C. Barrett (Ed.), *Teaching reading in the middle class* (pp. 51-58). Boston, MA.: Addison-Wesley.
- Brandão, A. C. P. & Oakhill, J. (2005). "How do you know this answer?" – Children's use of text data and general knowledge in story comprehension. *Reading and Writing*, 18, 687-713.
- Cadime, I. (2012). *Avaliar leitura no 1.º ciclo do ensino básico: Construção e validação do TCL-Teste de Compreensão Leitora*. Tese de doutoramento, não publicada. Braga: Universidade do Minho.
- Cain, K. & Oakhill, J. (1996). The nature of the relationship between comprehension skill and the ability to tell a story. *British Journal of Developmental Psychology*, 14, 187-201.
- Cain, K. & Oakhill, J. (2004). Reading comprehension difficulties. In T. Nunes & P. Bryant (Eds.), *Handbook of children's literacy* (pp. 313-338). London: Kluwer Academic Press.
- Cain, K., Oakhill, J. (2008). *Children's comprehension problems in oral and written language: A cognitive perspective*. New York: Guilford Press.
- Català, G., Català, M., Molina, E., & Monclús, R. (2001). *Evaluación de la comprensión lectora: Pruebas ACL (1º - 6º de primaria)*. Barcelona: Editorial Graó.
- Coelho, C. L. G. & Correa, J. (2010). Desenvolvimento da compreensão leitora através do monitoramento da leitura. *Psicologia: Reflexão e Crítica*, 23(3), 575-581.
- Colomer, T. & Camps, A. (2002). *Ensinar a ler, ensinar a compreender*. Porto Alegre: Artmed.
- Coscarelli, C. V. (2003). *Inferência: Afinal o que é isso?* Belo Horizonte: FAE/UFMG. Acedido em 9 fevereiro 2010: <http://bbs.metalink.com.br/~1coscarelli/publica.htm>.
- Cunningham, J. (1987). Toward pedagogy of inferential comprehension and creative response. In R. Tierney & J. Mitchell (Eds.), *Understanding reader's understanding* (pp. 229-255). New Jersey: Lawrence Erlbaum.
- Defior, S. (2000). *Dificuldades de aprendizagem: Un enfoque cognitivo. Lectura, Escritura, Matemáticas*. Granada: Ediciones Aljibe.
- Dell'Isola, R. L. P. (2001). *Leitura: Inferências e contexto sociocultural*. Belo Horizonte: Formato Editorial.
- Ehri, L. (1992). Reconceptualizing the development of sight word reading and its relationship to recording. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107-143). Hillsdale: Lawrence Erlbaum Associates.
- Fletcher, C. R., van der Broek, P., & Arthur, E. J. (1996). A model of narrative comprehension and recall. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 141-163). Mahwah: Lawrence Erlbaum.
- Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.
- Giasson, J. (2005). *La lecture: De la théorie à la pratique*. Bruxelles: De Boeck.

- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371-395.
- Graesser, A. C., Swamer, S. S., Baggett, W. B., & Sell, M. A. (1996). New models of deep comprehension. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 1-32). New Jersey: Lawrence Erlbaum Associates.
- Graesser, A. & Zwaan, R. (1995). Inference generation and the construction of situation models. In C. Weaver, S. Mannes & C. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 117-139). Hillsdale, NJ.: Lawrence Erlbaum Associates.
- Herber, H. (1978). *Teaching reading in content areas* (2nd Ed.). Englewood Cliffs, New Jersey: Prentice-Hall.
- King, A. (2007). Beyond literal comprehension: A strategy to promote deep understanding of text. In D.S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions and technologies* (pp. 267-290). New York: Lawrence Erlbaum Associates.
- Kintsch, W. (1998). *Comprehension, a paradigm for cognition*. Cambridge: Cambridge University Press.
- Kintsch, W. & Rawson, K. (2005). Comprehension. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 209-226). Oxford: Blackwell Publishing.
- Kleiman, A. (2002a). *Texto e leitor: Aspectos cognitivos da leitura*. Campinas: Pontes.
- Kleiman, A. (2002b). *Oficina de leitura: Teoria e prática*. Campinas: Pontes.
- Koch, I. V. & Elias, V. M. (2006). *Ler e compreender: Os sentidos do texto*. São Paulo: Editora Contexto.
- Marcuschi, L. A. (2008). *Produção textual, análise de gêneros e compreensão*. São Paulo: Parábola Editorial.
- Nation, K. (2005). Children's reading comprehension difficulties. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 248-265). Oxford: Blackwell Publishers.
- Oakhill, J. & Yuill, N. (1996). Higher order factors in comprehension disability: Processes and remediation. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 69-92). Mahwah, NJ: Lawrence Erlbaum.
- Pearson, P. D. & Johnson, D. D. (1978). *Teaching reading comprehension*. New York: Holt, Rinehart and Winston.
- Perfetti, C. A. (1985). *Reading ability*. Oxford: Oxford University Press.
- Perfetti, C. A., Marron, M., & Foltz, P. (1996). Sources of comprehension failure: Theoretical perspective and case studies. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 137-165). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Ruffman, T. (1996). Reassessing children's comprehension-monitoring skills. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 33-67). Mahwah, N. J: Lawrence Erlbaum Associates.

- Salles, J. F. & Parente, M. A. M. P. (2002). Processos cognitivos na leitura de palavras em crianças: relações com compreensão e tempo de leitura. *Psicologia: Reflexão e Crítica*, 15(2), 321-331.
- Salles, J. F. & Parente, M. A. M. P. (2004). Compreensão textual em alunos de segunda e terceira séries: Uma abordagem cognitiva. *Estudos de Psicologia*, 9(1), 71-80.
- Sim-Sim, I. (2001). *Avaliação da linguagem oral: Um contributo para o conhecimento do desenvolvimento linguístico das crianças portuguesas*. Lisboa: Fundação Calouste Gulbenkian.
- Solé, I. (1998). *Estratégias de leitura*. Porto Alegre: Artmed.
- Solé, I. (2003). Ler, leitura, compreensão: "Sempre falamos a mesma coisa?" In A. Teberosky, C. Oller, et al. (Eds.), *Compreensão de leitura: A língua como procedimento* (pp. 17-34). Porto Alegre: Artmed.
- Spear-Swerling, L., & Sternberg, R. J. (1996). *Off track: When poor readers became "learning disabled"*. Boulder, CO, US: Westview Press.
- Spinillo, A. G. (2013). A dimensão social, linguística e cognitiva da compreensão de textos: considerações teóricas e aplicadas. In M. M. P. E. Mota & A. G. Spinillo (Orgs.), *Compreensão de textos* (pp. 181-208). São Paulo: Casa do Psicólogo.
- Spinillo, A. G. & Mahon, E. R. (2007). Compreensão de texto em crianças: Comparações entre diferentes classes de inferência a partir de uma metodologia on-line. *Psicologia: Reflexão e Crítica*, 20(3), 463-471.
- Stanovich, K., West, R., Cunningham, A. E., Cipelewski, J., & Siddiqui, S. (1996). The role of inadequate print exposure as a determinant of reading comprehension problems. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 15-32). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Tolchinsky, L. & Pipkin, M. (2003). Seis leitores em busca de um texto. In A. Teberosky, et al. (Eds.), *Compreensão de leitura: A língua como procedimento* (pp. 95-105). Porto Alegre: Artmed.
- Vidal-Abarca, E. & Rico, G. M. (2003). Por que os textos são tão difíceis de compreender? As inferências são a resposta. In A. Teberosky, et al. (Eds.), *Compreensão de leitura: A língua como procedimento* (pp. 139-154). Porto Alegre: Artmed.
- Wechsler, D. (2003). *Escala de Inteligência de Wechsler para Crianças – Terceira edição (WISC – III)*. Adaptação portuguesa: M. R. Simões, A. Rocha & C. Ferreira. Lisboa: CEGOC-TEA.
- Yuill, N. & Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. Cambridge, MA: Cambridge University Press.

BATERIA DE AVALIAÇÃO DA LEITURA (BAL). DESCRIÇÃO E OPÇÕES DE CONSTRUÇÃO

Iolanda Ribeiro, Fernanda Leopoldina Viana, Sandra Santos, Irene Cadime, Alina Galvão Spinillo, Séli Chaves-Sousa e Ana Paula Vale

Neste capítulo oferece-se uma panorâmica geral dos testes que constituem a *Bateria de Avaliação da Leitura* (BAL), detalhando-se os objetivos, o público-alvo e as opções gerais de construção de cada um. A apresentação geral dos testes da BAL neste capítulo, apesar da existência de manuais específicos para cada um deles, deveu-se a duas razões.

Em *primeiro lugar*, os diferentes testes da BAL podem ser utilizados numa lógica de complementaridade, tendo-se clarificado no capítulo I deste livro a relevância de contemplar na avaliação das crianças com dificuldades de aprendizagem da leitura (DAL) as dimensões de leitura de palavras, de compreensão do que ouve (ler) e da compreensão da leitura silenciosa. No entanto, ressalva-se que, em função dos objetivos de cada utilizador, cada teste poderá ser aplicado individualmente e os seus resultados apreciados em função da informação que o mesmo fornece isoladamente. Note-se ainda que cada teste possui entre 3 a 4 versões, dirigidas a anos de escolaridade distintos. Mais uma vez, em função do objetivo da avaliação, o utilizador de cada teste pode optar por administrar apenas uma forma de teste, específica para determinado ano de escolaridade, ou no caso de se pretender avaliar os ganhos de um ou mais alunos ao longo do tempo, pode administrar as versões de teste correspondentes à medida que estes avançam na escolaridade;

Em *segundo lugar*, o facto de as opções metodológicas tomadas ao nível da construção dos testes de avaliação da compreensão de textos na modalidade de leitura e na modalidade oral serem comuns.

Assim sendo, considerou-se mais congruente clarificar neste volume as opções comuns a estes dois testes, nomeadamente as relativas à seleção dos textos, à construção dos itens e aos formatos de apresentação.

Plano de construção da BAL

A BAL é uma bateria de provas de avaliação da leitura que inclui cinco testes, organizados em torno de três domínios principais:

- 1) identificação de palavras apresentadas de forma isolada;
- 2) compreensão de textos na modalidade de leitura;
- 3) compreensão de textos na modalidade oral.

A designação de cada teste pode ser consultada no quadro 1.

Quadro 1. Apresentação dos testes que constituem a BAL.

DESIGNAÇÃO DO TESTE	AUTORAS	POPULAÇÃO-ALVO	VERSÕES DISPONÍVEIS
TLP Teste de leitura de palavras	Fernanda Leopoldina Viana Iolanda Ribeiro Ana Paula Vale Séli Chaves-Sousa Sandra Santos Irene Cadime	1.º ao 4.º ano do Ensino Básico	4
TCTML-n Teste de compreensão de textos na modalidade de leitura – narrativo	Iolanda Ribeiro Fernanda Leopoldina Viana Irene Cadime Sandra Santos Alina Galvão Spinillo	2.º ao 4.º ano do Ensino Básico	3
TCTML-i Teste de compreensão de textos na modalidade de leitura – informativo	Iolanda Ribeiro Fernanda Leopoldina Viana Irene Cadime Sandra Santos Alina Galvão Spinillo	2.º ao 4.º ano do Ensino Básico	3
TCTMO-n Teste de compreensão de textos na modalidade oral – narrativo	Iolanda Ribeiro Fernanda Leopoldina Viana Sandra Santos Irene Cadime	1.º ao 4.º ano do Ensino Básico	4
TCTMO-i Teste de compreensão de textos na modalidade oral – informativo	Iolanda Ribeiro Fernanda Leopoldina Viana Sandra Santos Irene Cadime	1.º ao 4.º ano do Ensino Básico	4

O desenvolvimento da BAL implicou a colaboração de um conjunto alargado de especialistas oriundos de diversas áreas de conhecimento. Foi, assim, possível uma abordagem multidisciplinar na elaboração dos testes. Contou-se, igualmente, com a colaboração de dois consultores externos especialistas em análise estatística – Gerardo Prieto (Universidade de Salamanca, Espanha) e José Maia (Faculdade de Desporto da Universidade do Porto) – na fase de análise de dados dos estudos de validação dos testes. Os textos que integram os testes de avaliação da compreensão foram elaborados por escritores consagrados na área da literatura infanto-juvenil: Ana Maria Magalhães, Carla Maia de Almeida, Isabel Alçada, Isabel Minhós Martins e Jorge Gomes. A revisão dos textos e dos itens dos testes ficou a cargo de especialistas do domínio da linguística – Armada Costa (Faculdade de Letras da Universidade de Lisboa), Rui Ramos (Instituto de Educação da Universidade do Minho) e José Morais (Universidade Livre de Bruxelas, Bélgica).

Construíram-se testes distintos para avaliar a compreensão de textos narrativos e informativos na modalidade oral (TCTMO-n e TCTMO-i) e na modalidade de leitura (TCTML-n e TCTML-i). Cada teste é composto por versões específicas para cada um dos anos de escolaridade que se visa avaliar. Os testes TLP, TCTMO-n e TCTMO-i são constituídos por quatro versões cada um, destinadas a alunos do 1.º ao 4.º ano de escolaridade.

Os testes TCTML-n e TCTML-i são constituídos por três versões cada, construídas para avaliar alunos do 2.º ao 4.º ano de escolaridade. No quadro 1 apresenta-se uma descrição geral de cada um dos testes que compõem a BAL.

Desenvolvimento dos testes da BAL: opções ao nível da construção e da administração

Avaliação da leitura de palavras: o TLP

O TLP (Teste de Leitura de Palavras) é um teste no qual os alunos devem ler em voz alta um conjunto de palavras, apresentadas isoladamente. Os itens (palavras) foram selecionados por peritos, de acordo com quatro critérios psicolinguísticos:

- a) extensão (curta; longa);
- b) frequência (muito frequente; pouco frequente);
- c) regularidade (regular; regra contextual; irregular);
- d) estrutura silábica (estruturas silábicas mais frequentes na Língua Portuguesa).

Na classificação da frequência das palavras recorreu-se à base de dados ESCOLEX (Soares *et al.*, 2013).

Foram construídas quatro versões do teste (uma para cada ano de escolaridade do 1.º ciclo), integrando 30 palavras cada. Cada versão de teste está disponível numa aplicação informática específica, desenvolvida para o efeito. As palavras que compõem cada versão de teste aparecem isoladamente no ecrã do computador, tendo a ordem de aparecimento sido aleatorizada. O aplicador do teste deve assinalar, durante a realização do mesmo, se cada palavra foi lida correta ou incorretamente. No final da administração do teste, a aplicação informática compila os resultados da avaliação num ficheiro Excel. Dado o seu formato, o TLP é um teste de administração individual.

A avaliação da compreensão: TCTML-n, TCTML-i, TCTMO-n e TCTMO-i

Procedimentos de medida

As medidas de compreensão podem ser classificadas como medidas em tempo real (*online*) ou em diferido (*offline*). Nas medidas em tempo real, a informação fornecida pela medida é obtida durante a leitura, enquanto nas medidas em diferido esta é recolhida após a realização da leitura (Lencastre, 2003). Os procedimentos de *think-aloud* e de registo dos movimentos sacádicos são exemplos de medidas pertencentes a este grupo. Para as medidas em diferido, o procedimento mais utilizado é a formulação de perguntas durante a leitura (i.e. a leitura é interrompida e o sujeito responde) ou no final da leitura (Lencastre, 2003). Os quatro testes de



compreensão que constituem a BAL podem classificar-se como medidas em diferido, dado que os itens são apresentados após a leitura (no caso dos testes correspondentes à modalidade de leitura) ou audição do texto (no caso dos testes correspondentes à modalidade oral).

Os textos

Na construção dos testes de compreensão optou-se por utilizar testes distintos para avaliar a compreensão de textos narrativos e de textos informativos, quer na modalidade oral, quer na modalidade de leitura. Os textos utilizados são originais e foram elaborados expressamente para o desenvolvimento da BAL, através de um trabalho de cooperação com os escritores, de modo a assegurar que os textos se ajustavam em termos de extensão e vocabulário aos níveis de escolaridade das crianças.

Foram adotadas as orientações de Sim-Sim & Viana (2007) na organização da estrutura e na extensão dos textos. Cada autor produziu entre dois a seis textos para os testes de compreensão:

- ◆ Ana Maria Magalhães e Isabel Alçada (2 textos no TCTML-n, 2 textos no TCTML-i e 2 textos no TCTMO-n);
- ◆ Jorge Gomes (2 textos no TCTML-i e 4 textos no TCTMO-i);
- ◆ Carla Maia de Almeida (1 texto no TCTML-n e 1 texto no TCTMO-n);
- ◆ Isabel Minhós Martins (1 texto no TCTML-n e 1 texto no TCTMO-n).

Os direitos de autor foram cedidos, tendo em vista a sua utilização exclusiva nos testes da BAL.

A opção por textos originais e não publicados anteriormente visou eliminar o efeito de um contacto prévio de alguns alunos com os textos, o que os colocaria em vantagem, relativamente ao seu grupo de referência. Em cada versão do TCTML-n e do TCTML-i foram incluídos 4 textos, de extensão crescente em função do ano de escolaridade (*cf.* quadro 2). A opção por utilizar os mesmos textos para todos os anos de escolaridade decorreu da necessidade de elaborar itens de ancoragem, isto é, itens comuns entre versões de teste para anos de escolaridade adjacentes. Os itens de ancoragem são fundamentais para a realização do processo de equalização de pontuações das diferentes formas de teste, tal como se explica no capítulo V. No entanto, e ainda que os textos sejam comuns aos três anos de escolaridade visados, a extensão destes é maior quanto mais avançado é o ano de escolaridade. Cada um dos textos incluiu o texto do ano anterior, pelo que a extensão é crescente. O texto de 3.º ano inclui o do 2.º e uma parte nova; o de 4.º ano inclui o de 3.º e, mais uma vez, é ampliado. A necessidade de manter os excertos nos vários textos é uma condição necessária para a construção de itens de ancoragem. Este plano de conceção dos textos permite avaliar os alunos usando um texto adequado à sua faixa etária/ano de escolaridade e diminuir as possibilidades de um efeito de teto na avaliação dos anos de escolaridade mais avançados.

Quadro 2. Organização dos textos nas versões do TCTML-n e TCTML-i.

TESTE	TEXTO	EXTENSÃO (nº de palavras)		
		2.º ano	3.º ano	4.º ano
TCTML-n Teste de compreensão de textos na modalidade de leitura – narrativo	A Rita e o Bruno	289	544	915
	Os animais	247	527	782
	Pão perdido	138	538	661
	Dois gémeos quase quase iguais	163	363	495
TCTML-i Teste de compreensão de textos na modalidade de leitura – informativo	Noites dos pirilampos	196	378	940
	As caravelas	179	351	838
	Observar aves no Estuário	173	450	718
	D. Sebastião	193	502	701

No TCTMO-n e TCTMO-i foram incluídos igualmente 4 textos em cada teste. Contrariamente ao TCTML-n e ao TCTML-i, os textos são comuns a todos os anos de escolaridade. No entanto, e para evitar uma sobrecarga da memória de trabalho, os textos foram segmentados, sendo a apresentação dos excertos alternada com a apresentação dos itens (cf. quadro 3).

Quadro 3. Organização dos textos nas versões do TCTMO-n e TCTMO-i.

TESTE	TEXTO	Nº DE PALAVRAS POR EXCERTO						
		1	2	3	4	5	6	7
TCTMO-n Teste de compreensão de textos na modalidade oral – narrativo	Um anjo na cozinha	86	149	135	123	84	-	-
	Férias na aldeia	109	195	180	117	130	122	-
	O filho do pirata	95	57	167	71	137	135	40
	Um gato com poderes especiais	117	142	123	174	106	-	-
TCTMO-i Teste de compreensão de textos na modalidade oral – informativo	Borboletas e lagartos	79	60	144	89	70	53	-
	Ai que solzinho!	120	136	154	90	130	57	-
	Ratinhos com asas	95	157	145	138	101	115	-
	Bugalhos e bolotas	141	110	125	100	158	-	-

A taxonomia adotada

Adotou-se a taxonomia de Barrett (1976), revista por Català, Català, Molina e Monclús (2001) na construção dos itens. Deste modo, cada item avalia a compreensão literal, a compreensão inferencial, a reorganização de informação ou a compreensão crítica.

Os itens

Nos quatro testes, os itens são constituídos por perguntas com resposta de escolha múltipla com três alternativas, das quais apenas uma é correta. Tem sido demonstrado que a inclusão de três alternativas nos

itens de escolha múltipla é, em termos psicométricos, tão eficiente como a inclusão de 4 ou 5 alternativas, tendo a vantagem de permitir uma aplicação mais rápida dos instrumentos de medida (Delgado & Prieto, 1998; Rodriguez, 2005). Para além de questões psicométricas, a inclusão de mais do que 3 alternativas poderia sobrecarregar a memória de trabalho nos testes de compreensão na modalidade oral. Por estes motivos, e de forma a manter a equivalência entre as duas modalidades de avaliação (leitura *versus* oral), optou-se por incluir 3 alternativas em todos os itens.

Quer os itens, quer as respetivas alternativas, foram sujeitos a revisão linguística por peritos. A posição em que cada alternativa correta é apresentada em cada um dos itens resultou de um processo prévio de aleatorização.

As modalidades: leitura *versus* oral

A opção por construir testes distintos para avaliar a compreensão nas modalidades de leitura e oral foi já discutida no capítulo I deste livro. Os dois tipos de testes distinguem-se também pelo formato de apresentação: os testes TCTML-n e TCTML-i são apresentados num formato de papel e lápis; os testes TCTMO-n e TCTMO-i em formato digital. Nos dois primeiros, tanto os textos como os itens são apresentados em papel. O aluno deve ler silenciosamente o texto e as perguntas e assinalar a resposta escolhida na folha de respostas. Estes testes podem ser aplicados individualmente ou em grupo.

Nos testes de compreensão de textos apresentados na modalidade oral, a leitura dos textos é ouvida com recurso a uma aplicação informática desenvolvida para o efeito, na qual são apresentados também os itens e gravadas as respostas dos alunos. Ressalve-se que, tanto o texto como os itens e respetivas alternativas de resposta são apresentados aos alunos apenas em formato áudio, sem a complementaridade de qualquer suporte escrito. A gravação da leitura dos textos e das perguntas foi efetuada por uma profissional. A administração dos dois testes pode ser feita individualmente ou em grupo, requerendo, neste último caso, a utilização de um computador e de auscultadores por criança. No final da administração do teste, a aplicação informática compila e exporta os resultados da avaliação para um ficheiro Excel.

As versões finais do TCTML-n são constituídas por 27 itens cada, incluindo-se 6 itens comuns (i.e. itens de ancoragem) entre versões adjacentes (22.2% do total de itens). Cada versão final do TCTML-i é composta por 33 itens, dos quais 8 (24.2%) são itens de ancoragem. As versões finais do TCTMO-n e TCTMO-i contêm 30 itens cada, nos quais se incluem 10 itens de ancoragem (33.3% do total das versões). Todas as versões dos testes possuem itens-exemplo.

Materiais que fazem parte da BAL

- ◆ O presente livro, composto por cinco capítulos, onde se incluem a descrição, os objetivos, as bases teóricas e as utilizações possíveis dos testes que a compõem.

TLP

- ◆ Manual técnico, onde se incluem os resultados dos estudos de construção e de validação do teste, os procedimentos para administração e cotação, bem como as respetivas orientações para interpretação de resultados;
- ◆ Quatro versões do teste em formato digital (uma por ano de escolaridade).

TCTML-n e TCTML-i

- ◆ Dois manuais técnicos (um para o TCTML-n e outro para o TCTML-i), onde se incluem os resultados dos estudos de construção e de validação dos testes, os procedimentos para administração e cotação, e as respetivas orientações para interpretação de resultados.
- ◆ Seis cadernos de teste (três para cada versão do TCTML-n e três para cada versão do TCTML-i).
- ◆ Seis folhas de resposta (três para cada versão do TCTML-n e três para cada versão do TCTML-i).

TCTMO-n e TCTMO-i

- ◆ Dois manuais técnicos (um para o TCTMO-n e outro para o TCTMO-i) onde se expõem os resultados dos estudos de construção e de validação dos testes, os procedimentos para administração e cotação, e as respetivas orientações para interpretação de resultados.
- ◆ Oito versões do teste em formato digital (quatro versões do TCTMO-n e quatro do TCTMO-i).

Referências bibliográficas

- Barrett, T. C. (1976). Taxonomy of reading comprehension. In R. S. T. C. Barrett (Ed.), *Teaching reading in the middle class* (pp. 51-58). Boston, MA.: Addison-Wesley.
- Català, G., Català, M., Molina, E., & Monclús, R. (2001). *Evaluación de la comprensión lectora: Pruebas ACL (1.º - 6.º de primaria)*. Barcelona: Editorial Graó.
- Delgado, A. R. & Prieto, G. (1998). Further evidence favoring three-option items in multiple-choice tests. *European Journal of Psychological Assessment*, 14(3), 197-201.
- Lencastre, L. (2003). *Leitura: A compreensão de textos*. Lisboa: Fundação Calouste Gulbenkian.
- Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24(2), 3-13.
- Sim-Sim, I. & Viana, F. L. (2007). *Para a avaliação do desempenho de leitura*. Lisboa: Ministério da Educação - Gabinete de Estatística e Planeamento da Educação (GEPE).



Soares, A. P., Medeiros, J. C., Simões, A., Machado, J., Costa, A., Iriarte, A., Almeida, J. J., Pinheiro, A., & Comesaña, M. (2013). ESCOLEX: A grade-level lexical database from European Portuguese Elementary to Middle School textbooks. *Behavior Research Methods*, 46(1), 240-253.

OPÇÕES METODOLÓGICAS E ESTATÍSTICAS NA CONSTRUÇÃO DOS TESTES DA BATERIA DE AVALIAÇÃO DA LEITURA

Irene Cadime, Iolanda Ribeiro e Fernanda Leopoldina Viana

A *Bateria de Avaliação da Leitura* (BAL) tem como objetivo avaliar a identificação de palavras escritas (TLP), a compreensão da leitura (TCTML-n e TCTML-i) e da linguagem oral (TCTMO-n e TCTMO-i). Pretende-se que os testes que a compõem permitam não só descrever diferenças interindividuais e comparar os resultados individuais dos leitores com os dos seus grupos de referência em termos de escolaridade, mas que, em simultâneo, permitam uma comparação intraindividual das mudanças de desempenho que ocorrem ao longo do tempo. Estas questões e a sua relevância foram analisadas no primeiro capítulo deste manual.

Quer a identificação das palavras escritas, quer a compreensão da leitura variam em função das aprendizagens dos sujeitos, sendo expectável que se observem ganhos ao longo da escolaridade (Bowey, 2005). Uma avaliação objetiva dos ganhos obtidos por cada aluno em diferentes momentos ao longo do tempo implica que as pontuações das medidas utilizadas em cada momento de avaliação sejam comparáveis. Esta comparação só é possível se for utilizada a mesma medida ou se as pontuações de diferentes medidas se encontrarem numa escala métrica comum. A opção pela utilização de uma mesma medida em vários momentos de avaliação ao longo da escolaridade é problemática. Para os alunos em anos iniciais de escolaridade, os testes poderão ser demasiado difíceis e, em contrapartida, corre-se o risco de encontrar efeito de teto nos anos de escolaridade mais avançados. Acresce que a aplicação sucessiva de uma mesma prova pode conduzir a um efeito de reatividade das medidas. A opção pela utilização de medidas diferentes para distintos níveis escolares é, neste caso, mais apropriada, pois tal permitirá que sejam aplicadas medidas com conteúdo e dificuldade ajustados ao nível de escolaridade em causa. No entanto, para que as pontuações de diferentes medidas sejam comparáveis, estas devem ser sujeitas a uma transformação que as coloque numa única escala, mediante um procedimento estatístico designado equalização (De Ayala, 2009). Embora existam diferentes alternativas para a realização da equalização, na BAL ela foi efetuada no quadro da Teoria de Resposta ao Item¹.

Todos os testes da BAL foram desenvolvidos com recurso às análises do modelo Rasch, tendo sido construídas versões específicas para cada ano de escolaridade. Após o estudo dos itens, as pontuações de cada versão de teste foram equalizadas, ou seja, foram colocadas numa mesma escala, de forma a que os utilizadores de cada teste possam comparar a evolução no desempenho de alunos avaliados ao longo do 1.º ciclo do Ensino Básico ou comparar as pontuações obtidas por alunos de anos de escolaridade distintos.

A fim de fornecer uma visão geral sobre as opções metodológicas e estatísticas na construção da BAL, neste capítulo começa-se por explicitar quais as limitações da Teoria Clássica dos Testes (TCT) que a Teoria da

¹ Uma descrição aprofundada dos métodos estatísticos mais adequados para a equalização de testes pode ser consultada em Kolen e Brennan (2010).

Resposta ao Item (TRI) permite ultrapassar. De seguida, exploram-se os princípios básicos da TRI, incluindo-se na mesma secção uma caracterização das bases teóricas do modelo Rasch. Posteriormente, descrevem-se as análises estatísticas específicas realizadas, para verificar, com recurso às análises do modelo Rasch, as propriedades psicométricas dos testes que compõem a BAL. Excetuam-se as análises relacionadas com a validade de critério e parte das normas construídas para os testes. Tanto o estudo da validade de critério como a determinação de percentis foram efetuados com recurso aos procedimentos clássicos. Ainda no âmbito das análises de TRI utilizadas, apresenta-se uma descrição da aplicação específica do modelo Rasch na equalização vertical de testes, com especial enfoque no método utilizado para a equalização das versões de cada um dos testes que integram a BAL.

Teoria de Resposta ao Item *versus* Teoria Clássica dos Testes

O desenvolvimento de testes ao longo da maior parte do século XX foi marcado pela utilização de um conjunto de procedimentos e análises estatísticas que são hoje identificados com uma corrente da psicometria denominada *Teoria Clássica dos Testes* (TCT) (Embretson & Reise, 2000). Embora não seja possível precisar a origem da TCT, esta afirmou-se como o grande referencial na psicometria, sobretudo a partir dos anos 30 do mesmo século, com a disseminação da utilização de testes estandardizados em vários contextos. As bases e os princípios da TCT são hoje sobejamente conhecidos. Uma descrição completa dos princípios da TCT pode ser consultada em Kline (2005). Apesar de ter sido usada de modo sistemático na construção de testes e provas psicológicas, a TCT possui várias limitações. O problema central da TCT foi desde cedo identificado, com o próprio Thurstone (citado por Pasquali & Primi, 2003) a assinalar, ainda antes dos anos 30, que “*um instrumento de medida, na sua função de medir, não pode ser seriamente afetado pelo objeto de medida. Na extensão em que sua função de medir for assim afetada, a validade do instrumento é prejudicada ou limitada. Se um metro mede diferentemente pelo fato de estar medindo um tapete, uma pintura ou um pedaço de papel, então nesta mesma extensão a confiança neste metro como instrumento de medida é prejudicada*” (pp. 99-100). Estamos, portanto, perante uma questão central da psicometria: a invariância das medidas. Este problema é frequentemente ilustrado, recorrendo-se ao exemplo dos testes de inteligência: se o grau de inteligência de um mesmo sujeito for avaliado com dois testes de inteligência distintos, os seus resultados, apesar de traduzirem ambos o nível do sujeito na competência, não são comparáveis. A solução adotada na TCT foi a de expressar as pontuações em função de parâmetros relativos definidos em função de um grupo normativo. Embora esta solução se constitua como útil para a prática, dificilmente é uma solução cientificamente aceitável, dado que inevitavelmente se define a variável medida em função do teste que a mede. No exemplo apresentado, a inteligência seria definida em função do teste aplicado ao sujeito, sendo a sua pontuação mais ou menos elevada em função do conjunto de itens utilizado (*test-dependent*). Tal pressuposto seria equivalente a dizer-se no campo da Física que o comprimento de um objeto depende da régua utilizada para o medir (Muñiz & Hambleton, 1992).

Outro problema central da TCT é o da invariância das propriedades dos instrumentos de medida. Na TCT a maioria das propriedades de um teste, como o índice de dificuldade ou o poder discriminativo, dependem do tipo de sujeitos utilizados para o seu cálculo (*group-dependent*). Centrando-nos por exemplo na dificuldade do teste, se um conjunto de itens for aplicado a uma amostra de sujeitos com alto nível de aptidão na variável medida, a dificuldade atribuída aos itens é menor do que se este for aplicado a um grupo de sujeitos com baixo nível de aptidão. Estabelecendo novamente uma analogia com o campo da Física, é como se as qualidades métricas de uma balança dependessem do tipo de objetos pesados (Muñiz & Hambleton, 1992).

Deste modo, na TCT existe uma dependência circular: a estatística dos sujeitos, ou seja, a pontuação que estes obtêm no teste é dependente do conjunto de itens utilizado na avaliação e as estatísticas dos itens, i.e. o índice de dificuldade e o poder discriminativo, são dependentes da amostra.

A resposta a estas limitações da TCT surgiu no quadro de um conjunto de modelos de traço latente, geralmente referidos como modelos de *Teoria de Resposta ao Item* (TRI). A TRI baseia-se no princípio geral de que o nível de traço latente ou aptidão dos sujeitos (variáveis não observáveis) pode ser estimado a partir de variáveis observáveis, tais como as respostas aos itens (Pasquali & Primi, 2003). Os modelos de TRI especificam como o nível de traço latente e as propriedades dos itens se relacionam com as respostas dos sujeitos a esses mesmos itens (Embretson & Reise, 2000). Estes modelos psicométricos oferecem uma solução para o problema da dependência da amostra e do teste, dado que possuem propriedades de invariância.

Tal como já se referiu, na TCT a pontuação do mesmo sujeito será mais alta se os itens forem fáceis e mais baixa se os itens forem difíceis, dado que o nível do sujeito na variável avaliada é dependente do teste com o qual esta é medida (*test-dependent*). Na TRI o cálculo da pontuação dos sujeitos é independente do conjunto de itens utilizado (*not test-dependent*) e, por isso, desde que os itens meçam o mesmo traço latente irão produzir o mesmo resultado em termos de aptidão do sujeito (Pasquali & Primi, 2003). Na TCT os parâmetros dos itens dependem, em larga medida, da amostra utilizada para o seu cálculo (*group-dependent*), mas na TRI o cálculo dos parâmetros dos itens é independente da amostra utilizada (*not group-dependent*), pelo que não é afetado pelas características da amostra. A propriedade de invariância dos parâmetros dos itens e da aptidão dos sujeitos é, assim, a característica central que diferencia a TRI da TCT, ao supor que os parâmetros que caracterizam o item não dependem da distribuição da aptidão dos sujeitos e que os parâmetros que caracterizam os sujeitos não dependem do conjunto particular de itens utilizado na sua avaliação (Arias, 1996).

Uma segunda limitação da TCT prende-se com a conceptualização do erro de medida. O erro padrão de medida refere-se à variação nas pontuações devido ao erro (Embretson & Reise, 2000). Relembre-se que, de acordo com os princípios básicos da TCT, a pontuação bruta e observável de um sujeito num determinado teste pode ser decomposta em duas partes: (a) a pontuação verdadeira, ou seja, o nível real do sujeito na variável medida; (b) o erro aleatório de medida, que ocorre em função da falta de precisão da medida utilizada. Na TCT o erro padrão de medida é obtido calculando-se a raiz quadrada de 1 menos o valor de fidelidade e multiplicando-a pelo desvio-padrão do teste (Embretson & Reise, 2000). Partindo deste cálculo, são construídos intervalos de confiança para as pontuações individuais, assumindo-se que o erro de medida está distribuído

normalmente e igualmente para todas as pontuações. Assume-se, assim, que há igualdade de variância ao longo de todo o contínuo de possíveis pontuações obtidas com um determinado teste (Pasquali & Primi, 2003). Esta suposição é bastante improvável, dado que a consistência das respostas está relacionada com o nível de aptidão dos sujeitos na competência avaliada (Hambleton & Swaminathan, 1985). Na TRI este problema não se coloca, uma vez que o pressuposto de igualdade de variância dos erros de medida não é assumido (Pasquali & Primi, 2003). Mais do que não assumir este pressuposto, largamente inverosímil, os modelos da TRI permitem quantificar o erro padrão da medida em cada nível da dimensão avaliada, possibilitando, por isso, a seleção de itens que permitem diminuir o grau de erro em regiões do atributo previamente especificadas (Prieto & Delgado, 2003).

A quarta limitação diz respeito ao cálculo da fidelidade ou da precisão que, na TCT, coloca também alguns problemas na construção de testes (Pasquali, 2003). Na TCT, as técnicas estatísticas para avaliar a fidelidade dos testes podem integrar-se em dois grupos principais: coeficientes de correlação e indicadores de consistência interna (Pasquali, 2009). A correlação é utilizada no método teste-reteste com o mesmo teste ou com formas paralelas (Arias, 1996). O método teste-reteste com a mesma forma de teste possui uma limitação evidente que torna o seu uso pouco recomendável, particularmente em testes de desempenho – o efeito de treino – dado que os participantes são testados duas vezes com o mesmo conjunto de itens. Também a correlação entre as pontuações obtidas em duas formas paralelas de um teste coloca o problema adicional da sua equivalência. Para que as formas de teste sejam estritamente paralelas, estas devem produzir variâncias semelhantes (Pasquali & Primi, 2003), o que na prática é extremamente difícil de obter. Os coeficientes de consistência interna como, por exemplo, o *alpha* de Cronbach, são uma alternativa para ultrapassar as limitações inerentes aos métodos assentes na correlação, mas possuem outras limitações. No caso do *alpha*, uma das críticas mais apontadas é a de que a sua fórmula tende a subavaliar a fidelidade da medida (Arias, 1996; Maroco & Garcia-Marques, 2006). Na TRI não existe a exigência de trabalhar com formas paralelas. Vários dos modelos que a integram possuem fórmulas de cálculo de coeficientes que estimam o grau de replicabilidade das posições dos sujeitos, se estes fossem sujeitos a versões paralelas de um teste, bem como o grau de replicabilidade das posições dos itens, se estes fossem estudados com uma amostra equivalente. Na exposição das análises a efetuar na construção de testes com recurso ao modelo Rasch, detalhar-se-ão dois desses índices – *person separation reliability* e o *item separation reliability*.

Teoria de Resposta ao Item: o caso específico do modelo Rasch

A TRI possui uma base teórica mais robusta do que a TCT e modela a distribuição probabilística de sucesso dos sujeitos a nível do item. A TRI foca-se primariamente a nível da informação do item, enquanto a informação primordial da TCT deriva da pontuação no total do teste. Dois postulados básicos regem a TRI: (a) o desempenho num teste pode ser predito a partir de um conjunto de aptidões ou traços latentes (geralmente designado pela letra grega *theta* [θ]), sendo estes a causa e a resposta aos itens o efeito;

(b) a relação entre o desempenho num item e a variável latente pode ser descrita por uma relação matemática monotónica crescente designada Curva Característica do Item, na qual a probabilidade de uma resposta correta ao item aumenta à medida que aumenta o *theta* do sujeito (Muñiz, 1996; Pasquali & Primi, 2003).

A TRI engloba uma diversidade de modelos cuja aplicabilidade a cada situação depende da natureza e da viabilidade de diferentes assunções teóricas sobre os itens do teste (Fan, 1998). Face às características da BAL, neste capítulo referir-se-ão apenas os modelos aplicáveis a itens dicotómicos e a situações onde o constructo a medir é unidimensional. Nos itens dicotómicos, as respostas dos sujeitos são cotadas como sucesso (1) ou insucesso (0). Apesar de marcados por diferenças na consideração de parâmetros para a modelação probabilística de respostas, duas assunções comuns subjazem a este conjunto de modelos: a unidimensionalidade e a independência local dos itens. Nos modelos unidimensionais, um único constructo latente é assumido como suficiente para caracterizar as diferenças entre os sujeitos, sendo a causa da resposta que cada sujeito dá ao item (Embretson & Reise, 2000). É este constructo latente que é suposto estar a ser medido pelo teste.

A independência local significa que a resposta a um item não pode estar relacionada com a resposta a qualquer outro item, quando o nível de aptidão é controlado (Muñiz, 1996). Na prática, a independência local é violada, quando as respostas entre os itens se encontram interligadas, como acontece nos casos em que o conteúdo de um item fornece informação relevante para a resposta correta a outro item. Itens que contêm terminologia muito relacionada ou que se referem a um estímulo comum, tal como os itens de testes de compreensão, podem ser particularmente suscetíveis à presença de dependência local (Embretson & Reise, 2000). Deste modo, a elaboração dos itens deve ser feita cuidadosamente de modo a garantir que em cada item não são fornecidas pistas para a resposta a itens subsequentes. Tanto a unidimensionalidade como a independência local dos itens podem ser testadas empiricamente, havendo várias propostas de métodos estatísticos adequados para esta testagem.

Neste sentido, pode ser considerada a utilização de três modelos logísticos: o modelo de um parâmetro, o de dois parâmetros e o de três parâmetros. Os três modelos assumem a função logística como função matemática que relaciona os valores da aptidão dos sujeitos na variável medida com a probabilidade de acerto a um item, representada graficamente pela curva característica do item (Muñiz & Hambleton, 1992). A distinção entre os três grupos de modelos é feita pelo número de parâmetros dos itens que cada modelo assume no cálculo desta probabilidade. Os modelos de um parâmetro têm em conta somente o parâmetro de dificuldade, enquanto os modelos de dois parâmetros recorrem à dificuldade e à discriminação do item. Os modelos de três parâmetros incluem, para além dos anteriores, a probabilidade de uma resposta correta ser fornecida ao acaso (Muñiz, 1996; Pasquali, 2003).

Ao longo deste capítulo, focar-nos-emos apenas no modelo que subjaz à construção dos diferentes testes que compõem a BAL: o modelo de um parâmetro, conhecido também por modelo Rasch, nome do seu principal impulsionador – Georg Rasch. Neste modelo a probabilidade da resposta correta a um item é definida em função do nível de aptidão do sujeito e de um parâmetro do item – a dificuldade (Bond & Fox, 2007). O modelo

Rasch é, assim, o mais parcimonioso ao caracterizar cada item em termos de um único parâmetro. O conceito de dificuldade aqui utilizado não é equivalente ao conceito de índice de dificuldade tradicionalmente associado à TCT, na qual este expressa a proporção de acertos em função do total de respondentes. No modelo Rasch, a dificuldade do item é um valor que indica a sua localização no *continuum* latente que representa o constructo (De Ayala, 2009). Na calibração dos resultados de um teste, recorrendo às análises do modelo Rasch, obtêm-se dois tipos de estimativas: um valor de aptidão para cada sujeito testado e um valor de dificuldade para cada um dos itens.

Estudo das propriedades psicométricas dos testes que integram a BAL

No estudo dos diferentes testes da BAL, contemplaram-se as linhas orientadoras para o desenvolvimento de testes de avaliação enunciadas no *Standards for educational and psychological testing* (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1999). Neste sentido, serão apresentados: (a) dados relativos ao desenvolvimento dos testes, especificando-se o modelo estatístico utilizado; (b) dados relativos à fidelidade dos testes; (c) dados relativos à normatização, comparabilidade de pontuações e equalização; e (d) evidência de validade.

Na construção e desenvolvimento dos testes da BAL recorreu-se às análises do modelo Rasch. O modelo possui um conjunto de análises que substitui as tradicionais análises da TCT de índice de dificuldade, poder discriminativo dos itens e consistência interna das medidas. Assim, mediante os procedimentos estatísticos do modelo Rasch, avaliaram-se as seguintes propriedades psicométricas dos itens e dos resultados dos diferentes testes que constituem a BAL:

- 1) parâmetros de itens (dificuldade) e sujeitos (aptidão);
- 2) estatísticas de ajuste;
- 3) unidimensionalidade, independência local dos itens;
- 4) fidelidade;
- 5) normatização, comparabilidade de pontuações e equalização;
- 6) evidência de validade.

A utilização do modelo Rasch na normatização dos testes, nomeadamente no processo de equalização, permite comparar as pontuações entre as versões de cada teste. Na normatização dos testes, calcularam-se também percentis para cada versão, de modo a permitir comparações por referência ao grupo normativo. No que respeita à evidência de validade, estudaram-se a validade de constructo e a validade de critério, recorrendo-se às análises clássicas para o estudo da validade de critério.

A apresentação dos dados psicométricos de cada teste nos manuais técnicos é efetuada de acordo com a sequência de análises apresentada neste capítulo.

Parâmetros dos itens (dificuldade) e dos sujeitos (aptidão)

O modelo Rasch possui a propriedade de medição conjunta, o que significa que os parâmetros de aptidão dos sujeitos e de dificuldade dos itens são expressos numa mesma escala, utilizando-se as mesmas unidades *Logit* (Bond & Fox, 2007). Esta propriedade tem a potencialidade de permitir analisar a adequação da dificuldade dos itens ao nível de competência dos sujeitos, bem como de inferir as interações entre sujeitos e itens. A interpretação das pontuações individuais pode, assim, basear-se, não só em normas de grupo, mas também na identificação dos itens a que cada sujeito tem alta ou baixa probabilidade de responder corretamente.

Tomemos como exemplo um teste hipotético onde o constructo a medir é a compreensão da leitura. Embora a compreensão da leitura não possa ser diretamente observada, é possível inferi-la a partir de observações comportamentais: a resposta aos itens do teste. Deste modo, a compreensão constitui-se como o constructo latente que, no modelo Rasch, é expressa num *continuum*. Os itens que constituem o teste são localizados em vários pontos do *continuum* que representa a compreensão da leitura. No mesmo *continuum* são localizados os diversos sujeitos da amostra. Deste modo, cada sujeito é caracterizado por um nível de aptidão (parâmetro do sujeito), cada item por um nível de dificuldade (parâmetro do item), e ambos são expressos por números ao longo de uma escala. A partir da diferença de valores entre os dois parâmetros, é possível calcular a probabilidade de uma resposta correta para cada sujeito a um dado item (Bond & Fox, 2007).

Na figura 2 (cf. página 70) apresenta-se um exemplo de uma representação gráfica da localização de itens e sujeitos no *continuum* de um hipotético teste de compreensão com sete itens (representação gráfica construída com recurso ao software WINSTEPS). Na figura 2, pode observar-se que os itens se encontram representados na parte direita do *continuum* e os sujeitos na parte esquerda.

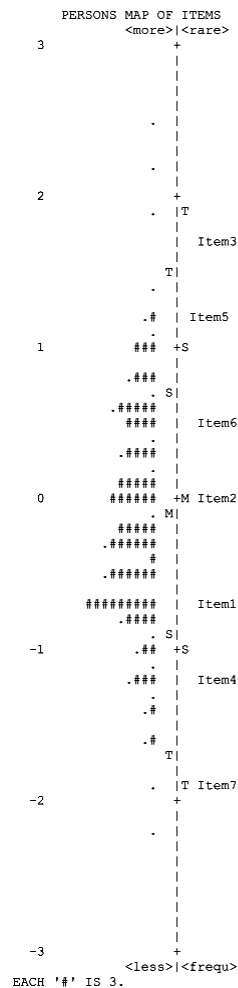


Figura 2. Representação gráfica da localização dos itens num *continuum* Rasch.

No caso específico da figura 2, o *continuum* varia entre -3 e 3 e a média dos itens está situada no valor zero. Ressalve-se, no entanto, que no modelo Rasch o intervalo de valores que o *continuum* assume não é estipulado *à priori*, mas determinado pelas estimativas da aptidão dos sujeitos e da dificuldade dos itens detetadas na matriz de dados (Bond & Fox, 2007). Para resolver a questão da indeterminação da escala, é necessária a definição de um ponto de referência, sendo geralmente adotada a solução de situar a média dos itens no valor zero. No entanto, é possível definir qualquer outro valor para a média dos itens ou, em alternativa, para a média dos valores paramétricos dos sujeitos. Por conseguinte, os valores da escala do *continuum* podem variar entre $-\infty$ e $+\infty$, embora, na maioria dos casos em que a média dos itens é situada em zero, os valores do *continuum* variem entre -5 e +5 (Prieto & Delgado, 2003).

Os valores paramétricos para a dificuldade do item e para a aptidão do sujeito são, assim, representados numa mesma escala logarítmica (expressa em valores *logit*), sendo atribuído a cada sujeito da amostra um valor de

aptidão e a cada item um valor de dificuldade. Tendo em conta ser assumido que valores mais elevados no *continuum* indicam maior nível de compreensão da leitura, quanto mais elevado for o parâmetro de localização do item, mais difícil este será e quanto mais elevada for a localização dos sujeitos, maior será a aptidão na variável medida. Observando a figura 2, verifica-se que o item mais difícil é o número 3 e o item mais fácil é o número 7. Verifica-se ainda que existem sujeitos distribuídos ao longo de todo o *continuum*, representando uma amostra heterogénea em termos de aptidão na compreensão da leitura. A relativa confluência das médias dos valores paramétricos de sujeitos e de itens indica que a dificuldade do teste é, em termos médios, adequada ao grupo de sujeitos avaliado.

Partindo dos valores paramétricos de sujeitos e itens, é possível calcular a probabilidade de um dado sujeito com um nível de aptidão x responder corretamente ao item y . A probabilidade de um sujeito responder corretamente a um item que esteja localizado exatamente ao mesmo nível do seu valor de aptidão é de 50%. Se observarmos a figura 2, a probabilidade de os sujeitos com um valor paramétrico de 0.00 responderem corretamente ao item 2 é de 50%, dado que se encontram ao mesmo nível no *continuum*. Itens com valores paramétricos muito distantes dos valores paramétricos dos sujeitos que se deseja avaliar, i.e. itens demasiado fáceis (aos quais, teoricamente, todos os sujeitos respondem), ou itens demasiado difíceis (aos quais nenhum sujeito consegue responder corretamente), são pouco informativos. Por conseguinte, deve ser ponderada a inclusão de tais itens, face aos custos que decorrem da aplicação de um teste mais longo em detrimento de um mais curto.

Na construção dos testes da BAL foram calculados os parâmetros de itens e sujeitos e foi avaliada a adequabilidade da dificuldade e do conteúdo dos itens para cada grupo, definido em termos de ano escolar. Em função dos resultados obtidos selecionaram-se os itens mais adequados para cada grupo, agrupando-os em versões específicas para cada ano escolar.

Estatísticas de ajuste: infit e outfit

Para verificar que o modelo Rasch é adequado para calibrar os dados, é indispensável que os mesmos se ajustem ao modelo, ou seja, que não haja diferenças significativas entre os resultados prognosticados pelo modelo e os verificados empiricamente. Se os dados se ajustarem ao modelo, a probabilidade de resposta correta a um dado item dependerá única e exclusivamente da aptidão do sujeito e da dificuldade do item. A presença de respostas anómalas, como, por exemplo, a constatação de sujeitos com alto nível de aptidão a falharem na resposta a itens muito fáceis ou sujeitos com baixo nível de aptidão a resolverem corretamente itens demasiado difíceis, indica que os dados empíricos carecem de ajuste ao modelo e que os valores para os parâmetros de itens e de sujeitos carecem de significado teórico. Embora não seja fácil determinar a causa do desajuste, tal pode dever-se a uma variedade de fatores: multidimensionalidade do teste, problemas ao nível do texto, da formulação dos itens ou das alternativas de resposta (nos casos em que existam), respostas ao acaso, entre outros².

² Para uma listagem mais completa de possíveis causas de desajuste consultar Karabatsos (2000).

O modelo Rasch possui indicadores de ajuste baseados em análises de resíduos – os indicadores de *Infit* e de *Outfit* (Bond & Fox, 2007). Estes dois índices fornecem informação relativa às discrepâncias nas respostas, consoante o seu afastamento dos parâmetros estimados. Os *softwares* para análises Rasch mais comumente utilizados, tais como o WINSTEPS, calculam estas estatísticas de ajuste tanto para os sujeitos como para os itens (de Ayala, 2008). Assim, a cada sujeito e a cada item é atribuído um valor de *Infit* e outro de *Outfit*.

O *Infit* é uma estatística ponderada baseada nos resíduos quadráticos estandardizados resultantes da diferença entre o que é observado e o que seria esperado com base no modelo (de Ayala, 2008; Prieto & Delgado, 2003). Para calcular o *Infit*, cada valor de resíduo quadrático no padrão de respostas, isto é, o resíduo “z” para cada um dos itens respondidos pelo sujeito “n”, é ponderado com a sua variância e somado. Este total é dividido pela soma das variâncias. Assim, o índice de *Infit* mais não é do que a média dos resíduos quadráticos, ponderada com a sua variância. Por sua vez, o *Outfit* é baseado na soma convencional dos resíduos quadráticos estandardizados, sem qualquer ponderação (Bond & Fox, 2007). Ambos os índices são reportados como médias quadráticas na forma de estatísticas de qui-quadrado divididas pelos seus graus de liberdade, de forma a situar o índice numa escala que pode variar entre 0 e $+\infty$, com um valor esperado de 1 (Bond & Fox, 2007; Prieto & Delgado, 2003). Os valores de *Infit* e de *Outfit* reportados como médias quadráticas não assumem, assim, valores negativos.

Um valor de *Infit* ou de *Outfit* de $1+x$ indica $100 \cdot x\%$ mais variação entre os dados observados e os padrões de resposta preditos pelo modelo do que a que seria esperada se os dados e o modelo fossem perfeitamente compatíveis (Bond & Fox, 2007). Deste modo, um valor de, por exemplo, 1.30 (i.e. $1+0.30=1.30$) indica 30% ($100 \cdot 0.30$) mais variação nos dados observados do que o estimado pelo modelo Rasch. Tomando outro exemplo com um valor inferior ao esperado, um valor de 0.78 ($1-0.22=0.78$) indica 22% ($100 \cdot 0.22$) menos de variação no padrão de respostas observado do que o previsto.

Os valores de *Infit* e de *Outfit* devem, idealmente, situar-se entre 0.5 e 1.5, não devendo ser superiores a 2.00, sob pena de haver distorção e degradação das medidas (Linacre, 2002). Itens com valores de desajuste superiores ao valor de referência 2.0 devem ser excluídos dos instrumentos. Valores inferiores a 0.5 não degradam as medidas e, portanto, não são preocupantes, mas podem ser indicadores de falhas a nível da independência local dos itens (Bond & Fox, 2007). Se o pressuposto da independência local não estiver assegurado, o padrão de respostas torna-se demasiado determinado, havendo muito pouca variabilidade não prevista pelo modelo, tornando redundante a informação fornecida pela medida.

Embora os valores de referência apresentados sejam indicados tanto para o índice ponderado como para o não ponderado, deve ressaltar-se que o *Outfit* é muito afetado pela presença de *outliers* ou respostas extremamente anómalas, como, por exemplo, respostas corretas a itens extremamente difíceis, por parte de sujeitos com muito baixo nível de aptidão (de Ayala, 2008; Pina & Montesinos, 2005). Tal facto foi também tido em conta na apreciação dos resultados das estatísticas de ajuste dos testes da BAL.

Unidimensionalidade e independência local dos itens

No programa estatístico WINSTEPS, tanto a unidimensionalidade como a independência local dos itens são testadas mediante análises de resíduos. Quando os dados são calibrados por meio do modelo Rasch, há observações que não se ajustam de forma perfeita ao modelo. Por conseguinte, destas diferenças entre o padrão de resultados observado e o padrão de resultados predito pelo modelo, resultam resíduos que podem ser analisados para explorar variância comum. *Clusters* de resíduos que partilham uma grande quantidade de variância comum podem ser identificados através da análise de resíduos em componentes principais.

A análise de resíduos em componentes principais não deve ser interpretada da mesma forma que a análise fatorial. Enquanto na análise fatorial o objetivo é extrair a estrutura fatorial mais adequada para “acomodar” a variância e a covariância dos dados observados, na análise de resíduos em componentes principais o objetivo é confirmar a hipótese de que os resíduos são “ruído” aleatório, através da extração da componente que explica a maior quantidade possível de variância dos resíduos (Brown, 2006; Linacre & Wright, 2001). Os *clusters* de resíduos, também designados dimensões secundárias, poderão ser considerados relevantes se explicarem uma percentagem elevada de variância e se tiverem a força de, pelo menos, dois itens. Por conseguinte, *eigenvalues* inferiores a 2.0 para as dimensões secundárias significam que a variância partilhada entre os resíduos pode ser considerada “ruído aleatório”, fornecendo suporte empírico para a unidimensionalidade dos dados (Linacre, 2011).

A análise das intercorrelações entre os resíduos é efetuada para examinar a possível existência de dependência local dos itens. Na base desta análise está a hipótese de que resíduos altamente correlacionados podem ser indicadores de que o desempenho dos sujeitos num item não depende apenas do seu nível de aptidão, ou seja do seu *theta*, mas que pode estar a ser “contaminado” pela resposta a outro item. De acordo com a sugestão de Linacre (2011), coeficientes de correlação superiores a .70 entre os resíduos de dois itens podem ser indicadores da existência de dependência local entre os mesmos.

Esta análise não só permite verificar dois pressupostos básicos do modelo Rasch, com implicações para a apreciação dos itens e para a sua seleção, como constitui um indicador de validade de constructo. Neste caso, a estrutura fatorial que se testa é a de que existe uma estrutura unidimensional. No processo de construção de testes com recurso à TCT, a questão da estrutura fatorial associada à validade de constructo é usualmente apresentada numa secção dedicada à análise dos resultados do teste. Na TRI, o teste da unidimensionalidade baseado na análise de resíduos é apresentado na secção dedicada ao processo de construção e de análise dos itens.

Fidelidade

O modelo Rasch permite o cálculo de dois coeficientes de fidelidade alternativos aos coeficientes clássicos da TCT: *person separation reliability* (PSR) e *item separation reliability* (ISR).

O coeficiente PSR indica quão eficientemente o grupo de itens consegue discriminar os sujeitos no constructo medido. Por sua vez, o coeficiente de ISR é um indicador da probabilidade de reprodução do nível de dificuldade dos itens ao longo de diversas aplicações. Os valores destes coeficientes variam entre 0 e 1 e, quanto mais elevado é o valor, melhor a separação entre os itens e os sujeitos e mais precisa a medida (Wright & Stone, 1999). Usualmente, utilizam-se valores de referência semelhantes aos utilizados para avaliar a adequabilidade do valor do *alpha* de Cronbach, pelo que é recomendável que estes superem o valor mínimo de .70.

Normalização, comparabilidade de pontuações e equalização

Para cada um dos testes que compõem a BAL elaboraram-se dois tipos de normas: centradas na média e no desvio-padrão (pontuações estandardizadas) e centradas na frequência (percentis).

As pontuações estandardizadas são notas *t* derivadas de uma transformação linear das pontuações *logit* obtidas na equalização das versões de cada teste. Quando os dados se ajustam ao modelo, uma das vantagens do uso da TRI está relacionada com a possibilidade de comparar os desempenhos de diferentes pessoas ou da mesma pessoa em diferentes ocasiões, através de uma técnica metodológica e estatística denominada equalização (do inglês *equating*).

Quando são utilizados testes diferentes (ou diferentes formas de um mesmo teste) e se pretende comparar os resultados obtidos em cada um, é requerido que estes se encontrem numa mesma escala métrica. Para tal, é necessário que os dados recolhidos sejam sujeitos a um processo de ligação entre versões ou formas, permitindo a comparabilidade de resultados. A equalização é o procedimento que permite ajustar as classificações a uma mesma escala, de modo que diferentes testes ou formas/versões de um teste possam ser usados, alternadamente, com um mesmo grupo de sujeitos ou aplicados a amostras diferentes e, ainda assim, os resultados obtidos em cada momento ou com cada grupo continuarem a ser comparáveis (Kolen & Brennan, 2010). Por outras palavras, trata-se de estabelecer uma correspondência entre as pontuações de diferentes testes ou versões/formas de um teste, de modo que estas, uma vez realizada a equalização, possam ser consideradas equivalentes.

Existem dois tipos de equalização: horizontal ou vertical. A primeira é utilizada quando os grupos possuem aptidões idênticas e os testes equalizados são de dificuldade semelhante. Apenas neste caso, as versões do teste podem ser utilizadas indistintamente. Por sua vez, o processo de equalização vertical, também chamado escalonamento vertical (do inglês vertical *scaling*), é empregue quando os grupos de participantes diferem no nível de habilidade e os testes diferem no nível de dificuldade (Baker, 1984). Este tipo de equalização é o indicado quando se pretende comparar desempenhos em aptidões ou competências que se supõe desenvolverem-se ao longo do tempo ou aumentar ao longo da escolaridade, como é o caso das competências de leitura. Na equalização vertical, as formas de teste não podem ser utilizadas indistintamente com cada grupo, pois cada forma de teste é dirigida a um grupo específico, como, por exemplo, a alunos de um determinado ano de escolaridade, sendo adequada em termos de dificuldade e de conteúdo apenas a esse mesmo grupo.

Mediante a aplicação dos procedimentos estatísticos de equalização vertical, as pontuações de cada forma de teste são colocadas num única escala, permitindo avaliar a evolução no desempenho de um aluno à medida que avança em termos de ano de escolaridade. Sendo esta uma das principais utilizações da BAL – avaliar a evolução no desempenho em competências de leitura ao longo da escolaridade – procedeu-se à equalização vertical das formas de cada um dos testes que integram esta bateria (TLP, TCTMO-n, TCTMO-i, TCTML-n e TCTML-i).

Consideraram-se duas fases centrais no processo de equalização: fase de recolha de dados e fase da transformação de pontuações (De Ayala, 2009). Relativamente à primeira, há que considerar os *designs* utilizados. No caso da equalização vertical, o *design* mais comumente adotado é o de grupos não equivalentes com itens de ancoragem. Este *design* implica a avaliação de dois ou mais grupos, aos quais é aplicada uma forma de teste específica. Na BAL os grupos correspondem a cada um dos anos de escolaridade. Cada versão de teste por ano de escolaridade inclui um grupo de itens únicos e um grupo de itens comum entre versões para grupos adjacentes (e.g. anos de escolaridade consecutivos). Estes itens comuns, também designados itens de ancoragem, são cruciais para o processo de equalização. De acordo com as recomendações da literatura, os itens de ancoragem devem idealmente constituir-se como uma “mini-versão” do teste total, ou seja, devem medir o mesmo constructo, representar os conteúdos avaliados e os seus parâmetros de dificuldade devem distribuir-se ao longo do *continuum* resultante da calibração do teste total (De Ayala, 2009; Kolen & Brennan, 2010). Na seleção de itens de ancoragem devem ainda privilegiar-se os que se mostrem razoavelmente estáveis ao longo das versões adjacentes que estão a ser equalizadas, ou seja, aqueles cuja dificuldade é semelhante entre grupos avaliados (Huynh & Meyer, 2010; Yu & Popp, 2005). Por fim, deve ainda ter-se em conta que o número de itens de ancoragem não deve ser inferior a 20% do total de itens do teste (Kolen & Brennan, 2010). Todos estes critérios foram contemplados no estudo e na seleção dos itens de ancoragem para cada versão dos testes que integram a BAL (TLP, TCTMO-i, TCTMO-n, TCTML-i e TCTML-n).

A segunda fase – transformação de pontuações – envolve dois passos sequenciais:

- a) a calibração dos dados com o modelo selecionado (modelo Rasch), a fim de obter evidência de ajuste ao modelo e os parâmetros dos itens para o posterior alinhamento da métrica das diferentes formas de teste;
- b) a aplicação efetiva do método de equalização (De Ayala, 2009).

Estão atualmente disponíveis vários métodos para a realização da equalização de testes no âmbito da TRI. Por razões de parcimónia, referir-nos-emos apenas ao método utilizado na equalização dos testes da BAL – o método de calibração com parâmetros de itens fixos³. No método de equalização selecionado, a calibração de cada forma de teste é efetuada separada e sequencialmente. Uma versão de teste é selecionada como o teste de referência e os parâmetros dos itens de ancoragem resultantes da sua calibração são utilizados para realizar a calibração de outra forma de teste – a forma de teste alvo (De Ayala, 2009). Na estimação dos parâmetros da forma de teste alvo, os valores dos itens de ancoragem são fixados nos valores obtidos na

calibração da forma de teste de referência e, por isso, não são re-estimados. Não sendo re-estimados, estes valores são utilizados para estimar os parâmetros dos itens únicos, colocando assim a forma de teste “alvo” na mesma escala de medida do teste de referência (Kolen & Brennan, 2010).

Na equalização das versões dos testes da BAL, selecionou-se como teste de referência a versão para o ano de escolaridade mais baixo. Após a calibração da referida versão de teste, os valores paramétricos dos itens de ancoragem foram utilizados para fixar os valores dos mesmos itens na calibração da versão para o ano de escolaridade subsequente, colocando-as assim na mesma métrica. Este processo foi repetido para as versões subsequentes, até todas as pontuações das versões que compõem um mesmo teste estarem localizadas numa escala comum. Estabeleceu-se a média das pontuações estandardizadas de cada escala em 100 (DP=10).

Evidência de validade

O estudo da validade de constructo foi efetuado no quadro do modelo Rasch, baseado na análise de resíduos já descrita na secção “unidimensionalidade e independência local dos itens” deste capítulo.

Esta não é uma abordagem muito frequente no estudo da estrutura interna dos testes, mas proporciona indicadores que permitem igualmente apreciar a estrutura fatorial dos mesmos. Tanto no caso dos testes de compreensão, como do teste de leitura de palavras é esperado que o modelo de um fator se ajuste aos dados observados. O teste da unidimensionalidade dos testes deriva, quer da definição teórica dos constructos avaliados (ver capítulos II e III), quer da observação dos resultados da análise da dimensionalidade de testes que avaliam constructos semelhantes (e.g. Cadime *et al.*, 2013).

No que respeita à validade de critério, examinaram-se as correlações entre os cinco testes da BAL, bem como entre os resultados nestes e noutras provas de compreensão (TCL - Cadime, Ribeiro, & Viana, 2012), de reconhecimento de palavras (PRP - Viana & Ribeiro, 2010), de memória de trabalho (Prova de memória de dígitos da WISC-III - Wechsler, 2003), de vocabulário (Prova de vocabulário da WISC-III - Wechsler, 2003) e de fluência (O REI - Carvalho, 2010).

Apresenta-se, de seguida, uma breve descrição de cada uma das provas utilizadas no estudo da validade dos testes da BAL:

TCL-Teste de Compreensão da Leitura (Cadime, Ribeiro, & Viana, 2012)

O TCL é um teste de avaliação da compreensão da leitura que avalia a compreensão literal e inferencial, a reorganização da informação e a compreensão crítica. É destinado a alunos do 2.º ao 4.º ano de escolaridade, com uma versão de teste específica para cada ano de escolaridade. O TCL é composto por um texto elaborado por uma escritora de literatura infanto-juvenil. Este está estruturado como um diário e integra sequências de texto narrativo, informativo, poético e prescritivo. Em cada item é apresentada uma pergunta e quatro

³ Uma descrição detalhada dos diversos métodos pode ser consultada em Kolen e Brennan (2010).

alternativas de resposta, das quais apenas uma é correta. A construção do TCL foi efetuada no quadro da teoria da resposta ao item. As versões destinadas a cada ano foram sujeitas a um processo de equalização de modo a estabelecer uma escala métrica comum a todas, o que permite comparar a evolução nos desempenhos dos sujeitos ao nível da compreensão da leitura. Os itens apresentaram valores de *Infit* e de *Outfit* inferiores a 1.50. Os valores dos coeficientes de Kuder-Richardson, *person separation reliability* e *item separation reliability* variaram entre .70 e .98. Os resultados do estudo da validade de constructo, efetuado com recurso à análise fatorial confirmatória, indicaram que os itens saturam num único fator em cada uma das versões. Em termos de validade de critério, foram obtidas correlações moderadas entre os resultados no TCL e outros critérios externos (Cadime *et al.*, 2013).

PRP-Prova de Reconhecimento de Palavras (Viana & Ribeiro, 2010)

A PRP é uma prova de reconhecimento de palavras regulares, destinada a avaliar a velocidade e a precisão de leitura, junto de alunos do 1.º ao 4.º ano do 1.º Ciclo do Ensino Básico. É constituída por 3 de treino e 40 itens experimentais. Em cada item é apresentada uma imagem, seguida de 4 palavras, de entre as quais os sujeitos têm de selecionar a que nomeia a imagem respetiva. Quanto à extensão, 20 palavras são dissilábicas e 20 são trissilábicas. Relativamente ao número de vizinhas ortográficas, em 5 itens são apresentadas 3 vizinhas da palavra-alvo, em 16 itens são apresentadas 2 vizinhas das palavras-alvo, em 13 itens é apenas apresentada uma vizinha e em 6 itens apenas são apresentadas palavras com proximidade ortográfica. O coeficiente *alpha* de Cronbach da prova variou entre .84 e .95. Os resultados da análise fatorial confirmatória indicam que todos os itens saturam num único fator. Em termos de validade de critério, as correlações com a avaliação da leitura realizada pelos professores variaram entre .62 e .84.

O REI-Teste de Avaliação da Fluência e Precisão da Leitura (Carvalho, 2010)

O REI é um teste de avaliação da fluência e da precisão da leitura, dirigido a alunos do 2.º ao 6.º ano de escolaridade. É constituído por um texto narrativo, com 281 palavras, que a criança deve ler em voz alta. O teste tem um tempo limite de 3 minutos para todos os anos de escolaridade. Ao longo da leitura são anotados os erros e é anotado o tempo total de leitura, caso a criança termine a leitura do texto antes de terminados os 3 minutos. O teste possibilita o cálculo de um índice de precisão e de um índice de fluência. O primeiro índice traduz a percentagem de palavras lidas corretamente e o segundo indica o número de palavras que a criança leu, em média, por minuto. Para efeitos do estudo de validade da BAL apenas se utilizou o índice de fluência. O teste apresenta bons indicadores de fidelidade, avaliada com recurso a teste-reteste, com correlações de .797 para o índice de precisão e de .938 para o índice de fluência. Em termos de validade, o teste apresenta correlações em geral moderadas com a avaliação da leitura efetuada pelos professores (Carvalho, 2010).

Memória de Dígitos (Wechsler, 2003)

Esta prova constitui um dos subtestes da Escala de Inteligência de Wechsler para crianças com idades compreendidas entre os 6 e os 16 anos – WISC-III – e constitui uma medida da atenção e da memória de



trabalho. São apresentadas várias séries de números à criança, incluindo uma como exemplo, e cada série é constituída por dois ensaios. A criança é solicitada a repetir cada série, contendo entre dois e nove dígitos, pela mesma ordem em que esta lhe foi apresentada (Dígitos em Sentido Direto) ou por ordem inversa (Dígitos em Sentido Inverso). A prova termina após insucesso em ambos os ensaios de uma mesma série. A pontuação atribuída corresponde a 1 ponto por cada série repetida corretamente. A cotação máxima para os Dígitos em Sentido Direto é de 16 pontos e para os Dígitos em Sentido Inverso é de 14 pontos. A WISC-III é uma medida que se encontra aferida para a população portuguesa e que apresenta boas qualidades psicométricas, revelando-se como um instrumento com elevada fidelidade e bons indicadores de validade.

Vocabulário (Wechsler, 2003)

A prova de vocabulário é outro dos subtestes da Escala de Inteligência de Wechsler para crianças - WISC-III. O subteste é constituído por um grupo de palavras apresentadas oralmente que o sujeito deve definir também oralmente. Cada resposta é cotada com 2, 1 ou 0 pontos. A aplicação do subteste é interrompida após 4 insucessos consecutivos. A cotação máxima é de 60 pontos.

No estudo de validade da BAL todos os testes da bateria, bem como o TCL e a PRP, foram administrados em grupo-turma, em contexto de sala de aula. O TLP, O REI, e os testes Memória de Dígitos e de Vocabulário foram aplicados individualmente.

Correlacionaram-se ainda os resultados nos testes da BAL com critérios externos de avaliação da descodificação e da compreensão pelos professores. A cada professor foi pedido que classificasse cada aluno em termos de:

- a) descodificação na leitura;
- b) compreensão na leitura;
- c) compreensão de textos lidos em voz alta.

Esta classificação foi efetuada com base numa escala de 1 (*Insuficiente*) a 5 (*Excelente*).

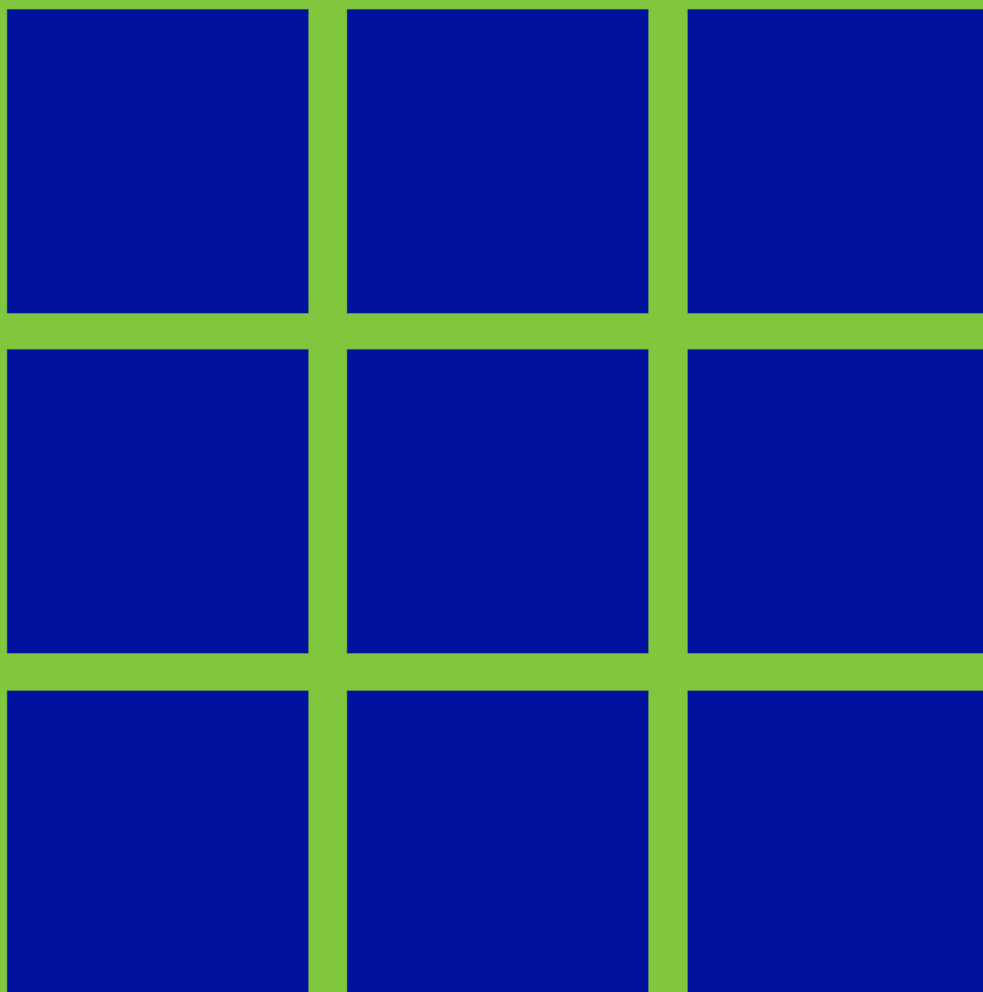
Referências bibliográficas

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Arias, R. M. (1996). *Psicometría: Teoría de los tests psicológicos y educativos*. Madrid: Editorial Síntesis.
- Baker, F. B. (1984). Ability metric transformations involved in vertical equating under item response theory. *Applied Psychological Measurement*, 8(3), 261-271.

- Bond, T. G. & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd Ed.). Mahwah, New Jersey: Lawrence Erlbaum.
- Bowey, J. (2005). Predicting individual differences in learning to read. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 155-172). Oxford: Blackwell Publishers.
- Brown, T. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.
- Cadime, I., Ribeiro, I., & Viana, F. L. (2012). *TCL - Teste de Compreensão da Leitura*. Lisboa: CEGOC-TEA.
- Cadime, I., Ribeiro, I., Viana, F. L., Santos, S., Prieto, G., & Maia, J. (2013). Validity of a reading comprehension test for Portuguese students. *Psicothema*, 25(3), 384-389. doi:10.7334/psicothema2012.288
- Carvalho, A. C. (2010). *Teste de avaliação da fluência e precisão da leitura: O REI*. Vila Nova de Gaia: Edipsico.
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York: The Guilford Press.
- Embretson, S. E. & Reise, S. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Fan, X. (1998). Item Response Theory and Classical Test Theory: An empirical comparison of their item/person statistics. *Educational and Psychological Measurement*, 58(3), 357-381. doi:10.1177/0013164498058003001
- Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory. Principles and applications*. Boston, MA: Kluwer Nijhoff Publishing.
- Huynh, H. & Meyer, P. (2010). Use of robust z in detecting unstable items in item response theory models. *Practical Assessment, Research & Evaluation*, 15(2), 1-8.
- Karabatsos, G. (2000). A critique of Rasch residual fit statistics. *Journal of Applied Measurement*, 1(2), 152-176.
- Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*. Thousand Oaks, CA: Sage Publications, Inc.
- Kolen, M. J. & Brennan, R. L. (2010). *Test equating, scaling and linking* (2nd ed.). New York: Springer.
- Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16(2), 878.
- Linacre, J. M. (2011). *A user's guide to WINSTEPS and MINISTEP: Rasch-model computer programs. Program manual 3.72.0*. Chicago, IL: Winsteps.
- Linacre, J. M. & Wright, B. D. (2001). *Winsteps* (Version 3.61.1) [Computer software]. Chicago: Mesa Press.
- Maroco, J. & Garcia-Marques, T. (2006). Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas? *Laboratório de Psicologia*, 4(1), 65-90.
- Muñiz, J. (1996). Fiabilidad. In J. Muñiz (Ed.), *Psicometría* (pp. 1-47). Madrid: Editorial Universitas.



- Muñiz, J. & Hambleton, R. K. (1992). Medio siglo de Teoria de Respuesta a los Ítems . *Anuario de Psicología*, 52, 41–66.
- Pasquali, L. (2003). *Psicometria: Teoria dos testes na psicologia e na educação*. Petrópolis: Editora Vozes.
- Pasquali, L. (2009). Psicometria. *Revista da escola de enfermagem da USP*, 43, 992-999.
- Pasquali, L. & Primi, R. (2003). Fundamentos da teoria da resposta ao item -TRI. *Avaliação Psicológica*, 2(2), 99-110.
- Pina, J. A. & Montesinos, M. D. (2005). Fitting Rasch model using appropriateness measure statistics. *The Spanish Journal of Psychology*, 8(1), 100-110.
- Prieto, G. & Delgado, A. R. (2003). Análisis de un test mediante el modelo de Rasch. *Psicothema*, 15(1), 94-100.
- Viana, F. L. & Ribeiro, I. (2010). *PRP - Prova de Reconhecimento de Palavras*. Lisboa: CEGOC-TEA.
- Wechsler, D. (2003). *Escala de Inteligência de Wechsler para Crianças – Terceira edição (WISC – III)*. Adaptação portuguesa: M. R. Simões, A. Rocha & C. Ferreira. Lisboa: CEGOC-TEA.
- Wright, B. D., & Stone, M. H. (1999). *Measurement essentials*. Wilmington, DE: Wide Range, Inc.
- Yu, C. H., & Popp, S. E. (2005). Test equating by common items and common subjects: concepts and applications. *Practical Assessment, Research & Evaluation*, 10(4), 1-19.



CEGOC - Testes Psicológicos
Rua General Firmino Miguel, n° 3, A/B, Rc
1600-100 Lisboa
Tel: 213 191 960
www.cegoc.pt